Eötvös Loránd University

Faculty of Science

Institute of Mathematics

Master's Thesis

Automorphic Forms and Expander Graphs

Author:
Zsolt Hoksza

Supervisor:
Gergely Harcos

Department of Analysis

2018

... à l'expansion de mon cœur refoulé s'ouvrirent aussitôt des espaces infinis. [1]

M. PROUST , À l'ombre des jeunes filles en fleurs

[1] According to my translation: ...the expansion of my repressed heart immediately opened up infinite spaces.

Acknowledgements

First and foremost, I wish to express my sincere thanks to my supervisor, Gergely Harcos, for introducing me into the beautiful theory of automorphic forms and proposing me such a topic which perfectly illustrates how these objects interact with other branches of mathematics. Our consultations had a fundamental contribution to understanding the basics of this theory and to my general professional development also. Besides, let me emphasize his permanent encouragement and enthusiasm which always motivated me and helped me to overcome the difficulties.

I would also like to extend my gratitude to Árpád Tóth, my consultant at Eötvös Loránd University. His inspiring thoughts and our discussions always helped me and broadened my perspectives. His lectures on complex analysis related topics were invaluable for my studies.

Finally, I would like to thank my parents and Eszter Kanti for their permanent support and encouragement.

Contents

Symbols

$\mathbb{Q}, \mathbb{Z}, \mathbb{R}, \mathbb{C}$	set of rational, integer, real and complex numbers
I	identity matrix or identity operator
$G = (V, E)$	graph on vertex set V with edges E
$\deg(G)$	$\max_{v \in V} \deg(v)$
$E(S, T)$	$\{(u, v) \in E \mid u \in S, v \in T\}$, $S, T \subseteq V$
∂S	edge boundary of S, $E(S, V \setminus S)$
$N(S)$	vertex boundary of S, $\{v \in V \setminus S \mid \exists u \in S \ (u, v) \in E\}$
$\chi(G)$	chromatic number of graph G
$\alpha(G)$	independence number of graph G
$\mathbb{T}_d$	the infinite d-regular tree
$h(G)$	expansion constant of G, $\min\left\{ \frac{\lvert\partial S\rvert}{\lvert S\rvert} \;\middle|\; S \subseteq V, \lvert S\rvert \leq \frac{1}{2}\lvert V\rvert \right\}$
$\mathfrak{S}_n$	the symmetric group of degree n
$\mathrm{Sp}(A)$	spectrum of operator A
r_A	spectral radius of operator A
$M_n(\mathbb{C})$	$\left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \;\middle|\; a, b, c, d \in \mathbb{C} \right\}$
$L^2(G, \mu)$	Hilbert space of square integrable functions on the vertex set with the usual inner product
A	adjacency operator of a graph, $Af(x) = \sum_{(x,y) \in E} f(y)$
$\mu(G)$	$\max(\lvert\mu_1\rvert, \lvert\mu_{n-1}\rvert)$ where $\mu_0 \geq \cdots \geq \mu_{n-1}$ are the eigenvalues of the adjacency operator
$\mu_1(G)$	the spectral gap of G, $\mu_0 - \mu_1$
χ_S	characteristic functions of set S
I	unit matrix
$p_k(v, w)$	number of paths of length k without backtracking from v to w
$p_k(x)$	$p_k(x, x)$
U_n	n-th Chebyshev polynomial of the second kind
δ_s	Dirac measure concentrated on $\{s\}$
$\hat{\mathcal{R}}$	universal covering space of a Riemann surface $\mathcal{R}$
$\pi_1(X)$	the fundamental group of a topological space X
$\left(\begin{smallmatrix} a & b \\ & d \end{smallmatrix}\right)$	$\left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right)$
$\left(\begin{smallmatrix} a & b \\ * & d \end{smallmatrix}\right)$	the asterisk denotes arbitrary matrix entry
$\mathrm{Aut}(X)$	set of (conformal) automorphisms of X
$GL_2(\mathbb{R})^+$	$\left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \;\middle|\; a, b, c, d \in \mathbb{R},\ ad - bc \neq 0 \right\}$
$SL_2(\mathbb{R})$	$\left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \;\middle|\; a, b, c, d \in \mathbb{R},\ ad - bc = 1 \right\}$
$PGL_2(\mathbb{R})$	$GL_2(\mathbb{R})/\{\pm I\}$
$PSL_2(\mathbb{R})$	$SL_2(\mathbb{R})/\{\pm I\}$
$j_\gamma(z)$	$j_\gamma(z) = (cz + d)$ if $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$
$\ell(\eta)$	length of the curve η
$\mathrm{Tr}(\gamma)$	the trace of the matrix γ
$\mathcal{F}_\Gamma$	the fundamental domain for Γ
Γ_z	stability subgroup, $\Gamma_z = \{\gamma \in \Gamma \mid \gamma z = z\}$
$\Gamma(q)$	$\left\{ \gamma \in SL_2(\mathbb{Z}) \;\middle|\; \left(\begin{smallmatrix} 1 & \\ & 1 \end{smallmatrix}\right) \pmod q \right\}$
$\Gamma_0(q)$	$\left\{ \gamma \in SL_2(\mathbb{Z}) \;\middle|\; \left(\begin{smallmatrix} * & * \\ & * \end{smallmatrix}\right) \pmod q \right\}$
$\Gamma_1(q)$	$\left\{ \gamma \in SL_2(\mathbb{Z}) \;\middle|\; \left(\begin{smallmatrix} 1 & * \\ & 1 \end{smallmatrix}\right) \pmod q \right\}$
$\nu(\gamma, z)$	automorphic factor
$\vartheta(\gamma)$	multiplier system
$f_{\mid\gamma}(z)$	slash operator of weight k, $\det(\gamma)^{k/2} j_\gamma(z)^{-k} f(\gamma z)$
$\mathfrak{a}$	an arbitrary cusp of the congruence subgroup Γ
$\sigma_\mathfrak{a}$	scaling matrix, $\sigma_\mathfrak{a} \infty = \mathfrak{a}$

$e(z)$	$e^{2\pi iz}$
$\theta(z)$	$\sum_{n\in\mathbb{Z}} e(n^2z)$
$r_Q(\nu)$	number of integer solutions of $Q(x)=\nu$, where Q is a quadratic form
$\theta_Q(z)$	$\sum_{\nu=0}^{\infty} r_Q(\nu)e(\nu z)$
$E_k^{(\infty)}(z)$	Eisenstein series, $\sum_{\gamma\in\Gamma_\infty} j_\gamma(z)^{-k}$
$\mathcal{M}(\Gamma,\vartheta)$	space of automorphic forms for Γ with multiplier system ϑ of weight k
$\mathcal{S}(\Gamma,\vartheta)$	space of cusp forms
$\mathcal{E}(\Gamma,\vartheta)$	space of Eisenstein series
T_n	n-th Hecke operator (except in Chapter 2)
$\tau(n)$	Ramanujan tau function
$\sigma_s(n)$	$\sum_{d\mid n} d^s$
$\zeta(s)$	Riemann zeta function
$\Delta(z)$	modular discriminant function
$\eta(z)$	Dedekind eta function
$\zeta(X,T)$	zeta function attached to variety X
$f_{\mid\gamma}(z)$	slash operator of weight k, $\det(\gamma)^{k/2}j_\gamma(z)^{-k}f(\gamma z)$
$\hat{G}$	universal cover of graph G (except in Chapter 5)
$\zeta_G(s)$	Ihara zeta function of graph G
$\tilde{H}$	Hurwitz quaternions
$H(\mathbb{Z})$	ring of integral Hurwitz quaternions
$\mathcal{G}(\Gamma,S)$	Cayley graph of group Γ with respect to $S\subseteq\Gamma$
T_pM	tangent space in p
$\mu_n(M)$	n-dimensional volume of a Riemannian manifold (M,g)
ω_g	volume form of an orientable Riemannian manifold (M,g)
Δ	Laplace-Beltrami operator
$e^{-t\Delta}$	heat operator
$h(M)$	Cheeger constant of the Riemannian manifold M
$\mathcal{U}(H)$	unitary group of a Hilbert space H
(π,H)	unitary representation
π_0	trivial representation
S^n	$\{x\in\mathbb{R}^{n+1} \mid \Vert x\Vert=1\}$
$SO_n(\mathbb{R})$	the group of orientation preserving isometries of $\mathbb{R}^n$
${}_gf$	$({}_gf)(x)=f(gx)$
$L_0^2(\Omega,\nu)$	$\left\{f\in L^2(\Omega,\nu) \;\middle|\; \int_\Omega f(\omega)\,\mathrm{d}\nu(\omega)=0\right\}$
$\varprojlim$	inverse limit

1. Introduction

The ancient Greeks asked: of all simple closed curves in the plane of a given length, which curve encloses the greatest area? Or equivalently, what is $\min\left(\ell(\partial N)^2/\mu_2(N)\right)$ if N is a region in $\mathbb{R}^2$ bounded by a simple closed curve $\partial(N)$, $\ell(\partial N)$ is the length of the curve, and $\mu_2(N)$ is the area of the region? The Greeks had no doubt about the correct answer to the classical question, namely the disk is the optimal shape in $\mathbb{R}^2$. The first rigorous proof of this question is due to Jacob Steiner (1841) based on the method called Steiner symmetrization.

The modern mathematicians twisted the question and asked what the optimal "shape" is on an n-dimensional Riemannian manifold M. The Cheeger constant of a compact n-dimensional Riemannian manifold M is $h(M) = \inf_N \left(\mu_{n-1}(\partial N)/\min(\mu_n(N), \mu_n(M \setminus N))\right)$, where N ranges over all n-dimensional submanifold of M whose complement $M \setminus N$ is also n-dimensional and has an $(n-1)$-dimensional boundary ∂N. It turns out that this geometric constant is strongly related to the value of the first non-trivial eigenvalue of a special operator of the Riemannian manifold, namely the Laplacian operator. The analogue of this question can be investigated in graph theory which leads us to the concept of expander graphs. This connection is presented in Chapter 5.

In Chapter 2 the concept of expander family is introduced. Expander graphs are highly connected sparse graphs which play an important role in applied and pure mathematics. To quantify these properties the expansion constant can be introduced. The expansion constant of a finite graph $G = (V, E)$ is defined by $h(G) = \min_S \frac{|\partial S|}{|S|}$ where S runs through the subsets of V with $|S| \leq \frac{1}{2}|V|$ and ∂S is the set of edges between S and its complement $V \setminus S$. Since every graph has a positive expansion constant, the notion is of interest only when one considers an infinite family of graphs whose expansion constants are bounded away from zero. A family of increasing graphs with uniformly bounded degrees is called expander family if it possesses the mentioned property.

These families can be characterized by spectral properties. If $G = (V, E)$ is a countable graph and μ is the counting measure, then one can consider the Hilbert space $L^2(G, \mu)$ of square-integrable complex valued functions on the vertex set V equipped with the usual inner product. Then the adjacency operator associated to the graph $A : L^2(G, \mu) \to L^2(G, \mu)$ can be defined by the formula

$$Af(x) = \sum_{(x,y)\in E} f(y).$$

Let $d = \mu_0 \geq \mu_1 \geq \ldots \mu_{n-1}$ be the eigenvalues of A. The spectrum of this operator is not only capable of encoding important combinatorial properties of the graph, but it is also suitable for spectrally characterizing the d-regular expansion families. A d-regular family of increasing graphs is an expander family if and only if the spectral gap $\mu_0 - \mu_1$ is bounded away from zero. This suggests the question how large the spectral gap of a d-regular finite graph can be.

The existence of expander families follows easily by random considerations but explicit constructions, which are very desirable for applications, are much more difficult. Deep mathematical theories have been used to give explicit constructions. The first explicit construction of an expander family of graphs is due to Margulis, who used Kazhdan's property (T) from representation theory of semi-simple Lie groups and their discrete subgroups. A group is said to have Kazhdan's property (T) if the trivial representation is an isolated point in its unitary dual equipped with the Fell topology. This means that if a unitary representation has almost invariant vectors, then it has a nonzero invariant vector. These concepts only have episodic role in this Thesis and are only introduced in Chapter 5.

Another construction due to Lubotzky, Phillips and Sarnak leads to a fascinating connection between the theory of expander graphs and the theory of automorphic forms via the Ramanujan conjecture for holomorphic cusp forms (proved by Deligne). Automorphic forms are among the most mysterious objects in mathematics. Amazing achievements of mathematics were born by successfully relating certain mathematical objects to automorphic forms, or more specifically to modular forms. For example the Modularity Theorem, or as formerly called the Shimura-Taniyama-Weil conjecture, states that all rational elliptic curves arise from modular forms. The theorem was proved for semistable elliptic curves by Wiles, completing the proof of Fermat's Last Theorem. The notion of modularity can be set in quite general context, but in this Thesis the classical description of them will be followed in Chapter 3. Let $\mathbb{H}$ be the upper half-plane $\mathbb{H} = \{z \in \mathbb{C} \mid \Im(z) > 0\}$. We denote by $SL_2(\mathbb{R})$ the group of 2×2 unimordular real matrices $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. This group acts on $\mathbb{H}$ via Möbius transformations: $z \to \frac{az+b}{cz+d}$. The functional equations defining modularity are of the type

$$f(\gamma z) = \nu(\gamma, z) f(z) \text{ for every } \gamma \in \Gamma \leq SL_2(\mathbb{R})$$

and for some simple and fixed function ν. If Γ is too large, then only the constants satisfy these functional equations, if Γ is the group of integer translations, then f is simply a periodic function. Let Γ be a suitable discrete subgroup of $SL_2(\mathbb{Z})$. The quotient space $\Gamma\backslash\mathbb{H}$ equipped with the topology in which the natural projection is continuous with properly chosen analytic charts becomes a Riemann surface. This surface can be compactified by adding cusps with suitable charts. A holomorphic function $f : \mathbb{H} \to \mathbb{C}$ is said to be a modular form if it satisfies the modularity functional equations and is holomorphic at the cusps of Γ. The meaning of this latter condition will be explained later. These functions have a Fourier expansion. One of the versions of the Ramanujan conjecture claims an upper bound for the Fourier coefficients of special modular forms, the so-called holomorphic cusp forms.

In Chapter 4 the construction due to Lubotzky, Phillips and Sarnak is presented. By their construction they introduced the notion of Ramanujan graphs. A finite, connected, d-regular graph is said to be Ramanujan if for every eigenvalue μ of the adjacency operator either $\mu = \pm d$ or $|\mu| \leq 2\sqrt{d-1}$. Considering the asymptotic results about eigenvalues and the spectral description of expander families, a family of Ramanujan graphs is a spectrally optimal expander family. The Ramanujan property of the constructed graphs follows from the Ramanujan conjecture for specific modular forms. The main purpose of this Thesis is to present this fascinating relation between the two theories.

2. Expander graphs

2.1 Family of expander graphs

Expander graphs, the subjects of this section, are certain growing families of sparse graphs with strong connectivity properties. The first step of our investigation will be to formalize the previous sentence. A family of finite graphs $(G_i)_{i\in I}$ is growing if the number of vertices $|V_i|$ goes to infinity, and the sparsity condition requires that the maximal degree of G_i is bounded by some constant. High connectivity is a natural strengthening of connectedness asking that any subset of vertices should have many connections with its complement.

2.1.1 Graph theory in general

Let us introduce some conventions now. A graph G is defined by a pair (V, E) where V is the set of vertices and $E \subset V \times V$ is the set of edges. Self loops and multiple edges are allowed. The degree of a vertex v is denoted by $\deg(v)$ and $\deg(G) = \max\{\deg(v) \mid v \in V\}$. A graph is called (n, d)-graph if it is finite with $n > 0$ vertices, connected and $\deg(v) = d$ for all $v \in V$. For $S, T \subseteq V$ denote the set of edges from S to T by $E(S, T) = \{(u, v) \in E \mid u \in S, v \in T\}$. The edge boundary of a set S denoted ∂S is $\partial S = E(S, V \setminus S)$. This is the set of edges emanating from the set S to its complement. The vertex boundary of a set S is defined by $N(S) = \{v \in V \setminus S \mid \exists u \in S \ (u, v) \in E\}$. A graph $G = (V, E)$ is bipartite if there exists a partition $V = V^{(1)} \cup V^{(2)}$ of the vertex set into two disjoint subsets so that any edge has one endpoint in $V^{(1)}$ and one in $V^{(2)}$. For any two vertices $v, w \in V$ the distance between v and w, denoted by $d(v, w)$ is defined as the minimum length of a path between v, w, if such a path exists, or $+\infty$ otherwise. A geodesic in G is a path p such that the length of p is equal to the distance between the endpoints of p. The distance function d is a metric on V. The diameter of a graph, denoted $\mathrm{diam}(G)$, is the largest distance between two vertices in G, i.e., $\mathrm{diam}(G) = \sup_{v,w\in V} d(v, w)$. The girth of the graph is the length of a shortest cycle contained in the graph and it is denoted by $\mathrm{girth}(G)$. The chromatic number $\chi(G)$ is the smallest integer $k \geq 0$ such that there exists a k-coloring of V, i.e., there is a function $c : V \to \{1, \dots, k\}$ such that $c(v) \neq c(w)$ whenever $(v, w) \in E$. The independence number $\alpha(G)$ is the largest integer k such that there exists $S \subseteq V$ with the property that elements of S are not connected. The finite tree of degree d and depth k, denoted $\mathbb{T}_{d,k}$, is a simple graph defined by taking

$$V = \bigcup_{0\leq j\leq k} \left\{(s_1, \dots, s_j) \in \mathcal{A}^j \mid s_i \neq s_{i+1} \text{ for } 1 \leq i \leq j - 1\right\}$$

where $\mathcal{A} = \{1, \dots, d\}$. The elements of V are called words and the elements of $\mathcal{A}$ are letters. The vertex set contains the empty word which is called the root vertex of the tree. Two vertices v and w are connected if and only if w can be obtained from v either by adding a letter on the right or by removing the last letter. The infinite d-regular tree $\mathbb{T}_d$ (for $d \geq 2$) is the infinite graph with vertices given by all words of length at least 0, without repeated letter, in the alphabet $\{1, \dots, d\}$, and with edges described above.

2.1.2 Isoperimetric problem on graphs

It is a natural choice to measure the area of a set $S \subseteq V$ in a graph by the cardinality of $|S|$ while a plausible measure for the boundary can be the number of edges exiting from the set. This motivates the introduction of the expansion constant of a graph.

Definition 2.1 (Expansion constant)**.** Let $G = (V, E)$ be a finite graph. The (edge) expansion constant $h(G)$ is defined by

$$h(G) = \min \left\{ \frac{|\partial S|}{|S|} \in [0, +\infty) \;\middle|\; \emptyset \neq S \subset V,\ |S| \leq \tfrac{1}{2}|V| \right\}$$

with the convention that $h(G) = +\infty$ if G has at most one vertex.

Proposition 2.2. *Let $G = (V, E)$ be a finite graph with at least two vertices.*

(i) $h(G) > 0$ *if and only if* G *is connected.*

(ii) *If* $S \subset V$ *and* $|S| = \delta|V|$ *where* $0 < \delta \leq \frac{1}{2}$*, one must remove at least* $\delta h(G)|V|$ *edges from* G *to disconnect* S *from the rest of the graph.*

(iii) $\frac{2}{|V|} \leq h(G) \leq \min\limits_{v \in V} \deg(v)$.

Proof. (i) The condition $h(G) = 0$ means that there exists a non-empty subset S of V, such that $|S| \leq \frac{1}{2}|V|$ and $|\partial S| = 0$. This immediately implies that the graph has at least two connected components. Conversely if G is not connected, let S be a connected component of minimal size in G. Since its size is minimal, we have $|S| \leq \frac{n}{2}$, and since S is a connected component, it is disconnected from the rest of the graph, i.e. $\partial S = \emptyset$. This set makes expansion coefficient equal to zero.

(ii) This follows directly from the definition. For all $S \subset V$ with $\delta|V| = |S| \leq \frac{1}{2}|V|$ we have that $|\partial S| \geq h(G)|S| = \delta h(G)|V|$. This condition can be translated as there are at least $h(G)|S|$ edges which go from S to $V \setminus S$. If one wants to disconnect the set S, then these edges have to be removed. In the light of this argument, the statement is just a reformulation of the definition.

(iii) Since G is connected for any non-empty proper subset $|\partial S| \geq 1$, we have $\frac{|\partial S|}{|S|} \geq \frac{1}{|S|} \geq \frac{2}{|V|}$. The upper bound can be seen by applying the expansion condition to the set $S = \{v\}$, where $v \in V$ is a vertex with minimal degree. □

Every finite connected graph with more than one vertex has a positive expander constant in a trivial way. The notion is of interest only when we consider an infinite family of graphs. This motivates the next definition.

Definition 2.3 (Expander family)**.** A family $(G_i)_{i \in I}$ of non-empty connected finite graphs $G_i = (V_i, E_i)$ is an expander family, if there exists constant $h > 0$, such that:

(i) For any $n \geq 1$, there are only finitely many $i \in I$ such that $|V_i| \leq n$.

(ii) $\sup_{i \in I} \deg(G_i) < \infty$.

(iii) For each $i \in I$, the expansion constant satisfies $h(G_i) \geq h$.

While the counting measure is an absolutely natural choice for a subset S of the vertex set V, the expansion constant could have been defined by using the number of extremities out of the particular subset, i.e., using $|N(S)|$, instead of the quantity $|\partial S|$. In the case of a bipartite graph, another variant of the expansion constant can be obtained by looking just at one of the partitioning subsets.

Definition 2.4 (Vertex-expansion constant)**.** Let $G = (V, E)$ be a finite graph. The vertex-expansion constant $h_v(G)$ is defined by

$$h_v(G) = \min\left\{ \frac{|N(S)|}{|S|} \in [0, +\infty) \;\middle|\; \emptyset \neq S \subset V,\ |S| \leq \tfrac{1}{2}|S| \right\}.$$

If G is a bipartite graph with a decomposition $V = V^{(1)} \cup V^{(2)}$, let

$$\hat{h}(G) = \min\left\{ \min\left\{ \frac{|N(S)|}{|S|} \in [0, +\infty) \;\middle|\; \emptyset \neq S \subset V^{(i)},\ |S| \leq \tfrac{1}{2}|S| \right\} \;\middle|\; i = 1, 2 \right\}$$

Lemma 2.5. *Let $G = (V, E)$ be a non-empty finite graph. If $S \subset V$ then we have*

$$\frac{1}{\deg(G)}|\partial S| \leq |N(S)| \leq |\partial S|$$

Proof. Consider the map $\varphi_S : \partial S \to V$ which sends an edge in ∂S to the one among its endpoints which is not in S, i.e. $\varphi_S(e) = e \cap (V \setminus S)$. By definition this map is surjective, which implies $|N(S)| \leq |\partial S|$. On the other hand, if $v \in N(S)$, then $|\varphi_S^{-1}(\{v\})| \leq \deg(G)$ since at most $\deg(G)$ edges have v as extremity. Summing this inequality over all $v \in N(S)$ we get $|\partial S| = |\varphi_S^{-1}(N(s))| \leq \deg(G)|N(S)|$ which was to be demonstrated. $\square$

Lemma 2.6. *Let $G = (V, E)$ be a finite bipartite connected graph with a decomposition $V = V^{(1)} \cup V^{(2)}$ where $|V^{(1)}| = |V^{(2)}|$. Then*

$$\frac{\hat{h}(G) - 1}{2} \leq h(G) \leq \deg(G)\hat{h}(G)$$

Proof. (Upper bound) Since the expansion constant is a minimum over a set, an upper bound means that the set has an element which is at least that large. Hence an equivalent reformulation of the inequality is the following: there exists $S \subset V : |S| \leq \frac{1}{2}|V|$ and $\frac{|\partial S|}{|S|} \leq \deg(G)\frac{|N(W)|}{|W|}$ for all $W \subset V^{(i)}$ and $i = 1, 2$. Let $w \in V$ be a vertex with $\deg(w) = \deg(G)$. Then

$$\min_{\substack{S \subset V \\ |S| \leq \frac{1}{2}|V|}} \frac{|\partial S|}{|S|} \leq \frac{\deg(G)}{|\{w\}|} \leq \deg(G) \leq \deg(G) \min\left\{ \min_{S_i \subset V^{(i)}} \frac{|N(S_i)|}{|S_i|} \;\middle|\; i = 1, 2 \right\}.$$

(Lower bound) Let $|V| = 2n = 2|V^{(1)}| = 2|V^{(2)}|$. As we saw in the previous argument, we may assume that $\hat{h}(G) \geq 1$ and hence we can rewrite it in the form $\hat{h}(G) = 1 + \delta$ with $\delta \geq 0$. In this case, the lower bound equivalent to the condition $h(G) \geq \frac{\delta}{2}$. Let $S \subset V$ be an arbitrary subset with $|S| \leq n$. Then S can be partitioned into $S^{(1)} \cup S^{(2)}$ where $S^{(j)} = S \cap V^{(j)}$ for $j = 1, 2$. We can assume that $|S^{(1)}| \leq |S^{(2)}|$ and hence $\frac{1}{2}|S| \leq |S^{(2)}|$. Two cases can be distinguished whether $|S^{(2)}| \leq \frac{1}{2}n$ or not.

In the first case, if $|S^{(2)}| \leq \frac{1}{2}n$ then we can apply the definition of the bipartite vertex-expansion which yields $|S^{(2)}|(1 + \delta) \leq |N(S^{(2)})|$ where $N(S^{(2)}) \subset V^{(1)}$. Among the neighbours of $S^{(2)}$, at most $|S^{(1)}| \leq |S^{(2)}|$ belong to S, i.e., $\left|N(S^{(2)}) \setminus S^{(1)}\right| \geq \delta|S^{(2)}| \geq \frac{\delta}{2}|S|$ hence $|N(S)| \geq \frac{\delta}{2}|S|$ and we are done.

In the second case, $|S^{(2)}| > \frac{1}{2}n$. The definition cannot be applied directly to $S^{(2)}$ because its size is too large, but can be applied to a subset of size $\lceil n/2 \rceil$. This implies that $\frac{1}{2}n(1+\delta) \leq |N(S^{(2)}|$. Then $\left|N(S^{(2)}) \setminus S^{(1)}\right| \geq (1 + \delta)\frac{n}{2} - \frac{n}{2} = \frac{\delta}{2}n = \frac{\delta}{4}|V| \geq \frac{\delta}{2}|S|$. $\square$

As a consequence of the proved inequalities, in the setting of expander families there is no difference between the classes of graphs distinguished by the variants of the expansion constants.

Corollary 2.7. *A family $(G_i)_{i\in I}$ of non-empty connected finite graphs $G_i = (V_i, E_i)$. Let us assume that for any $n \geq 1$, there are only finitely many $i \in I$ such that $|V_i| \leq n$ and $\sup_{i\in I} \deg(G_i) < \infty$. Then the following are equivalents:*

(i) *$(G_i)_{i\in I}$ is an expander family, i.e., there exists $c_1 > 0$ such that for each $i \in I$, $h(G_i) \geq c_1$.*

(ii) *There exists $c_2 > 0$ such that for each $i \in I$, $h_v(G_i) \geq c_2$.*

(iii) *There exists $c_3 > 0$ such that for each $i \in I$, $\hat{h}(G_i) \geq c_3$.*

The following statement shows that large expansion coefficient implies that the diameter of a graph is relatively small.

Proposition 2.8 (Expansion and diameter)**.** *Let G be a (n, d)-graph. Then the following inequality holds*

$$\operatorname{diam}(G) \leq 2\frac{\log \frac{n}{2}}{\log\left(1 + \frac{h(G)}{d}\right)} + 3.$$

The proof can be found in Kowalski's lecture notes on expander graphs [Kow17]. For further information on expander graphs one can also see [HLW06], [Lub12].

2.1.3 Existence of expander families

By a probabilistic argument below we will demonstrate the existence of expander graphs. In fact, it turns out that for many models of random graphs, there is a positive lower bound for the expansion constant which holds with high probability. Essentially the same proof can be found in many related books and articles ([Kow17] , [HLW06], [Sar90], [Lub10]).

Theorem 2.9 (Existence of expanders)**.** *Fix $d \geq 3$. Let $\sigma = (\sigma_1, \dots \sigma_d)$ be a d-tuple of random variables σ_i uniformly distributed on the set $\mathfrak{S}_n$ for each $1 \leq i \leq d$. For fixed $n \geq 1$ and any d-tuple σ of permutations of $\{1, \dots, n\}$, we define a graph $G_{n,\sigma}$ with vertex set*

$$V_{n,\sigma} = \big\{(j, 0) \mid 1 \leq j \leq n\big\} \cup \big\{(j, 1) \mid 1 \leq j \leq n\big\} = V^{(1)} \cup V^{(2)}$$

and edge set $E_{n,\sigma} = \big\{\big((j, 0), (\sigma_i(j), 1)\big) \mid 1 \leq j \leq n,\ 1 \leq i \leq d\big\}$ Then there exists $h_d > 0$ such that $\lim_{n\to\infty} \mathbb{P}\left(h(G_n, \sigma) < h_d\right) = 0$.

Proof. First let us note that in our model for random graphs, the probability that the random graph $G_{n,\sigma}$ possesses a given property can be calculated the following way:

$$\mathbb{P}\left(G_{n,\sigma} \text{ has property } \varphi\right) = \frac{1}{|\mathfrak{S}_n|^d}\big|\{\sigma \in \mathfrak{S}_n^d \mid G_{n,\sigma} \text{ has property } \varphi\}\big|.$$

We prove that $h_d = \frac{1}{2}$ is a suitable constant for $d \geq 5$. Since the graph $G_{n,\sigma}$ defined above is bipartite, Lemma 2.6 yields the inequality

$$\limsup_{n\to\infty} \mathbb{P}\left(h(G_\sigma) < h_d\right) \leq \limsup_{n\to\infty} \mathbb{P}\left(\hat{h}(G_\sigma) < 1 + 2h_d\right).$$

Hence it is sufficient to show that $\lim_{n\to\infty} \mathbb{P}\left(\hat{h}(G_{n,\sigma}) < 1 + 2h_d\right) = 0$. Let n be fixed. By symmetry, we have

$$\mathbb{P}\left(\hat{h}(G_{n,\sigma}) < 1+\delta\right) = \mathbb{P}\left(\min_{\substack{S\subset V^{(1)}\\ 1\le|S|\le\frac{n}{2}}} \frac{|N_\sigma(S)|}{|S|} < 1+\delta\right) \le \sum_{\substack{S\subset V^{(1)}\\ 1\le|S|\le\frac{n}{2}}} \mathbb{P}\left(|N_\sigma(S)| < (1+\delta)|S|\right).$$

Again, for symmetry reasons, the probability $\mathbb{P}\left(|N_\sigma(S)| < (1+\delta)|S|\right)$ depends only on $|S|$, because any subset of size t in $V^{(1)}$ is equivalent to $\{1,\dots,t\}$ in this model.

$$\sum_{\substack{S\subset V^{(1)}\\ 1\le|S|\le\frac{n}{2}}} \mathbb{P}\left(|N_\sigma(S)| < (1+\delta)|S|\right) \le \sum_{1\le t\le\frac{n}{2}} \binom{n}{t} \mathbb{P}\left(\left|N_\sigma(\{1,2,\dots t\})\right| < (1+\delta)t\right)$$

Since σ_i is a bijection for every i, and there are edges joining $(j,0)$ and $(\sigma_1(j),1)$ for all j, we always have $|N_\sigma(\{1,\dots,t\})| \ge t$. The condition $\left|N_\sigma(\{1,\dots,t\})\right| < (1+\delta)t$ implies that there is $H \subseteq V^{(2)}$ such that $t \le |H| \le (1+\delta)t$ and $N_\sigma(\{1,\dots,t\}) = H$. Or equivalently for $i = 1,\dots,d$, $\sigma_i(\{1,\dots,t\}) \subseteq H$. Hence

$$\mathbb{P}\left(|N_\sigma(\{1,2,\dots t\})| < (1+\delta)t\right) \le \sum_{\substack{H\subset V^{(2)}\\ |H|\le(1+\delta)t}} \mathbb{P}\left(\left|\sigma_i(\{1,2,\dots t\})\right| \subseteq H \;\middle|\; i=1,2,\dots d\right).$$

By definition of the random graph, the permutations $\{\sigma_i\}_{i=1}^d$ are independent. The probability of that for a particular i and a particular $H \subseteq V^{(2)}$ the condition $\sigma_i(\{1,\dots,t\}) \subseteq H$ holds is $m(m-1)\dots(m-t+1)(n-t)!$ which implies

$$\sum_{t=1}^{\lfloor n/2\rfloor} \binom{n}{t} \sum_{\substack{H\subseteq V^{(2)}\\ |H|\le(1+\delta)t}} \mathbb{P}\left(\sigma_i(\{1,\dots t\}) \subseteq H\right)^d = \frac{1}{(n!)^d} \sum_{t=1}^{\lfloor n/2\rfloor} \sum_{m=t}^{\lfloor(1+\delta)t\rfloor} \binom{n}{t}\binom{n}{m}\left(\frac{m!(n-t)!}{(m-t)!}\right)^d.$$

$\delta = \frac{1}{2}$ will be a suitable choice as the rest of the proof will demonstrate. Calculating with this constant, the sums in our upper bound are

$$\sum_{t=1}^{\lfloor n/2\rfloor} \sum_{m=t}^{\lfloor 3t/2\rfloor} \binom{n}{t}\binom{n}{m}\left(\frac{m!(n-t)!}{(m-t)!}\right)^d.$$

We split the exterior sum at $t = \lfloor n/3\rfloor$ because slightly different techniques are required to handle the different parts.

$$\sum_{t=1}^{\lfloor n/3\rfloor} \sum_{m=t}^{\lfloor 3t/2\rfloor} \binom{n}{t}\binom{n}{m}\left(\frac{m!(n-t)!}{(m-t)!}\right)^d + \sum_{\lceil t=n/3\rceil}^{\lfloor n/2\rfloor} \sum_{m=t}^{\lfloor 3t/2\rfloor} \binom{n}{t}\binom{n}{m}\left(\frac{m!(n-t)!}{(m-t)!}\right)^d$$

We write S_1 and S_2 to further examine these two quantities separately. Then using the bounds $t \le m \le \lfloor 3t/2\rfloor$ we have

$$S_1 \le n \sum_{t\le\lfloor n/3\rfloor} \binom{n}{t}\binom{n}{\lfloor\frac{3t}{2}\rfloor}\left(\frac{\lfloor\frac{3t}{2}\rfloor!(n-t)!}{\lfloor\frac{t}{2}\rfloor!}\right)^d.$$

Let us introduce the notation $R(t)$ for the terms in the sum. By checking $R(t)/R(t+1)$ we see that for $t < \lfloor\frac{n}{3}\rfloor$ this is at least 1, and so at this range $R(t)$ gets its maximum for $t = 1$.

$$S_1 \le n(\tfrac{n}{3})\binom{n}{1}^2\left((n-1)!\right)^d$$

If n goes to infinity then we also have

$$\lim_{n\to\infty} \frac{S_1}{(n!)^d} \le \lim_{n\to\infty} n^{4-d} = 0$$

for $d \ge 5$. On the other hand,

$$S_2 \le n\big(2^{2n}\big) \sum_{t=\lceil n/3\rceil}^{\lfloor n/2\rfloor} \left(\frac{\lfloor \frac{3t}{2}\rfloor!(n-t)!}{\lfloor \frac{t}{2}\rfloor!}\right)^d = n\big(2^{2n}\big) \sum_{t=\lceil n/3\rceil}^{\lfloor n/2\rfloor} R'(t).$$

It can be showed that $R'(t)$ achieves a maximum at one of the endpoints, either at $t = \lceil n/3\rceil$ or $t = \lfloor n/2\rfloor$. Checking at these points in combination with Stirling's formula one finds that

$$\lim_{n\to\infty} \mathbb{P}\left(\hat{h}(G_{n,\sigma}) < 1 + 2h_d\right) \le \lim_{n\to\infty} \left(\frac{S_1}{(n!)^d} + \frac{S_2}{(n!)^d}\right) = 0.$$

□

Although the probabilistic argument above shows the existence of expander families, an explicit construction turns out to be much more difficult.

2.2 Spectral theory of graphs

In this section a brief introduction to the spectral theory of graphs will be given in order to obtain an analytical description of the expander property. The vertex set of a graph can be equipped with measures. If one considers the functions on this set then the L^2 Hilbert space of functions can be defined in the usual way. The analysis of special linear operators acting on this Hilbert spaces provides us a powerful tool to investigate the combinatorial properties of graphs with the full apparatus of functional analysis.

2.2.1 Functional analysis in general

A complex normed vector space is called a Banach space if it is complete with respect to the metric induced by the norm. Let $(X, \|.\|_X)$ and $(Y, \|.\|_Y)$ be complex Banach spaces. A linear operator $A : X \to Y$ is called bounded if there exists a constant $c \ge 0$ such that $\|Ax\|_Y \le c\|x\|_X$ for all $x \in X$. The smallest constant $c \ge 0$ that satisfies this is called the operator norm of A, i.e., $\|A\| = \sup_{x\in X\setminus\{0\}} \frac{\|A\|_Y}{\|x\|_X}$.

$\mathcal{L}(X,Y)$ denotes the space of bounded complex linear operators from X to Y which is a complex Banach space with respect to the operator norm. In the case $X = Y$ abbreviate $\mathcal{L}(X) := \mathcal{L}(X,X)$. The dual space of a complex Banach space X is the space $X^* := \mathcal{L}(X,\mathbb{C})$ of bounded complex linear functionals. If X is a complex Banach space and $A \in \mathcal{L}(X)$ then let us define the spectrum of A as the set $\mathrm{Sp}(A) = \{\mu \in \mathbb{C} | \nexists (\mu\mathbf{1} - A)^{-1}\} = \mathrm{Sp}_p(A) \cup \mathrm{Sp}_r(A) \cup \mathrm{Sp}_c(A)$, where $\mathrm{Sp}_p(A)$ is the point spectrum, $\mathrm{Sp}_c(A)$ is the continuous spectrum and $\mathrm{Sp}_r(A)$ is the residual spectrum. These are defined by $\mathrm{Sp}_p(A) = \{\mu \in \mathbb{C} \mid (\mu\mathbf{1} - A)$ is not injective $\}$, $\mathrm{Sp}_c(A) = \{\mu \in \mathbb{C} \mid (\mu\mathbf{1} - A)$ is injective, its image is dense but it is not surjective $\}$ and finally $\mathrm{Sp}_r(A) = \{\mu \in \mathbb{C} \mid (\mu\mathbf{1} - A)$ is injective but its image is not dense$\}$. A complex number μ belongs to the point spectrum $\mathrm{Sp}_p(A)$ if and only if there exists a nonzero vector $x \in X$ such that $Ax = \mu x$. The elements $\mu \in \mathrm{Sp}_p(A)$ are called eigenvalues of A and the nonzero vectors $x \in \mathrm{Ker}(\mu\mathbf{1} - A)$

are called eigenvectors. The point spectrum will be the most important for us, because we will use this terminology mostly in context of adjacency matrices. In this case when $\dim X = n$ then $\mathrm{Sp}(A) = \mathrm{Sp}_p(A)$ and $|\,\mathrm{Sp}(A)| \leq n$. The spectral radius of a bounded linear operator $A : X \to X$ on a complex Banach space is the real number $r_A = \inf_{n\in\mathbb{N}} \|A^n\|^{1/n} = \lim_{n\to\infty} \|A^n\|^{1/n} \leq \|A\|$. The reason for the terminology spectral radius is the next theorem.

Theorem 2.10 (Properties of the spectrum)**.** *Let X be a nonzero complex Banach space and let $A \in \mathcal{L}(X)$.*

(i) $\mathrm{Sp}(A) \neq \emptyset$

(ii) $r_A = \lim\limits_{n\to\infty} \|A^n\|^{1/n} = \sup\limits_{\mu\in\mathrm{Sp}(A)} |\mu|$

(iii) $\mathrm{Sp}(A)$ *is a compact subset of* $\mathbb{C}$.

A complex Hilbert space $(H, \langle .,. \rangle)$ is a complex vector space H equipped with a Hermitian inner product such that the norm is complete. If X and Y are Hilbert spaces and $A \in \mathcal{L}(X, Y)$ then one can define the adjoint operator of A as the unique operator $A^* : Y \to X$ that satisfies the formula $\langle Ax, y\rangle_Y = \langle x,^* y\rangle_X$ for all $x \in X$ and all $y \in Y$.

Theorem 2.11 (Spectrum of self-adjoint operators)**.** *Let H be a nonzero complex Hilbert space. Let $A \in \mathcal{L}(H)$ be a self-adjoint operator, i.e., $A = A^*$. Then the following holds*

(i) $\mathrm{Sp}(A) \subset \mathbb{R}$.

(ii) $\|A\| = \sup\limits_{\|x\|=1} |\langle Ax, x\rangle| = \sup\limits_{x\neq 0} \dfrac{|\langle Ax, x\rangle|}{\langle x, x\rangle}$.

Theorem 2.12 (Spectral theorem)**.** *Let H be a finite-dimensional Hilbert space. Then the linear operator $A \in \mathcal{H}$ is self-adjoint if and only if H is the orthogonal direct sum of the eigenspaces of A for real eigenvalues.*

This is a fundamental theorem in mathematics. Since every symmetric matrix defines a self-adjoint operator on a finite-dimensional Hilbert space, we have the following corollary.

Corrollary 2.13. *Let $A \in M_n(\mathbb{R})$ be a real symmetric matrix. Then A is orthogonally equivalent to a diagonal matrix, i.e., $A = U^T DU$ with $U^T U = I$ and*

$$\begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix}$$

where $\mu_1, \ldots, \mu_n$ are the eigenvalues of A with multiplicity.

2.2.2 Spectrum of a graph

Definition 2.14. Let $G = (V, E)$ be a countable graph and μ be the counting measure, i.e., $\mu(\{x\}) = 1$ for $x \in V$. Then $L^2(G, \mu)$ is the Hilbert space of square-integrable complex valued functions on the graph, i.e.:

$$L^2(G,\mu) = \left(\left\{f : V \to \mathbb{C} \mid \int_V |f(x)|^2 \,\mathrm{d}\mu(x) < +\infty\right\}, \langle .,. \rangle\right)$$

where $\langle f, g\rangle = \int_V f(x)\overline{g(x)}d\mu(x) = \sum_{x\in V} f(x)\overline{g(x)}$ is the usual inner product.

Definition 2.15. Let $G = (V, E)$ be a countable graph with bounded degrees. Define the adjacency operator $A : L^2(G, \mu) \to L^2(G, \mu)$ by the formula

$$Af(x) = \sum_{(x,y)\in E} f(y)$$

In fact, $Af(x)$ is the sum of f over all of the neighbours of x. If G is finite graph and one enumerates the vertices of G as $v_1, v_2, \dots v_n$, then one can associate to A an $n \times n$ matrix, known as the adjacency matrix of G. As our graphs are undirected, the adjacency operator A is self-adjoint as its matrix is symmetric. By Corollary 2.13 A has n real eigenvalues with multiplicity. We denote the eigenvalues by $\mu_0 \geq \mu_1 \geq \cdots \geq \mu_{n-1}$. By the spectrum of G we mean the spectrum of the adjacency operator A. From the multiplication rule of matrices we see that if $A^k = \left(a_{ij}^{(k)}\right)$ then $a_{ij}^{(k)}$ is the number of walks in G of length k from v_i to v_j.

Proposition 2.16. *If $G = (V, E)$ is a finite d-regular graph on n vertices then*

(i) $d = \mu_0 \geq \mu_{n-1} \geq -d$;
(ii) *the multiplicity of μ_0 is the number of connected components;*
(iii) $\mu_{n-1} = -d$ *if and only if G is bipartite;*
(iv) *the eigenvalues are symmetric about* 0 *if and only if G is bipartite;*
(v) $\displaystyle\sum_{i=0}^{n-1} \mu_i^k = \operatorname{Tr}\left(A^k\right) = \left|\{(v_1, v_2, \dots v_m) \mid v_{i+1} \in N(v_i), m \leq k, v_1 = v_m\}\right|.$

The proof of this proposition is not really challenging. Any eigenvalue $\mu_j \neq \pm d$ is referred to as a non-trivial eigenvalue.

Definition 2.17 (Spectral parameters)**.** Let G be a d-regular graph on n vertices with eigenvalues $\mu_0 \geq \mu_1 \geq \cdots \geq \mu_{n-1}$. The spectral parameters of the graph are defined by the equations

$$\mu(G) = \max\left(|\mu_1|, |\mu_{n-1}|\right), \ \mu(G)^* = \max_{|\mu_j|\neq d} |\mu_j|, \quad \text{and } \mu_1(G) = \mu_0 - \mu_1.$$

The last quantity $\mu_1(G)$ is called the spectral gap of G.

If G is a d-regular graph Proposition 2.16 shows that $\mu(G) = \mu(G)^*$ unless G is a bipartite graph. In this case $\mu(G) = d$. The main reason of the examination of the spectrum for us is that the expander property of a graph has a hallmark in the spectrum as the following theorem shows.

Theorem 2.18 (Spectral gap theorem)**.** *Let $G = (V, E)$ be a (n, d)-graph, i.e., $|V| = n$, connected and d-regular. Let $\mu_1(G)$ be the spectral gap of G. Then*

$$\frac{\mu_1(G)}{2} \leq h(G) \leq \sqrt{2d\mu_1(G)}.$$

This was proved by Dodziuk and independently by Alon-Milman and by Alon. The proof can be found in [HLW06] or in [Sar90]. From this theorem it can be seen that the spectral gap provides an estimate on the expansion constant of a graph. Furthermore, we get the following spectral description of expander families.

Corrollary 2.19 (Spectral characterization of expander families)**.** *A family $(G_i)_{i\in I}$ of non-empty connected finite graphs $G_i = (V_i, E_i)$ is an expander family, if there exists constant $\varepsilon > 0$, such that:*

(i) *For any $n \geq 1$, there are only finitely many $i \in I$ such that $|V_i| \leq n$.*
(ii) $\sup_{i \in I} \deg(G_i) < \infty$.
(iii) *For each $i \in I$, the spectral gap satisfies $\mu_1(G_i) \geq \varepsilon$.*

A family of d-regular graphs is a family of expanders if and only if the spectral gap is bounded away from zero. This is the property that made possible for Margulis to use the apparatus of representation theory in order to construct expander graphs. Finally, we end this section by proving a trivial lower bound for $\mu(G)$.

Proposition 2.20. *Let G be an (n, d)-graph. Then $\mu(G) \geq \sqrt{d}\big(1 - o_n(1)\big)$.*

Proof. Since $\mathrm{Tr}(A^k)$ is the number of all walks of length k in G that start and end in the same vertex, the diagonal entries in A^2 are at least d. Hence $nd \leq \mathrm{Tr}(A^2) = \sum_{i=0}^{n-1} \mu_i^2 \leq d^2 + (n-1)\mu(G)^2$, and from this follows that $\mu \geq \sqrt{d}\big(\frac{n-d}{n-1}\big)^{1/2} = \sqrt{d}\big(1 - o_n(1)\big)$ as claimed. □

2.2.3 Combinatorial properties and the spectrum

The spectrum is not only capable of characterizing the expander property but encodes other combinatorial properties also. Let S and T be disjoint subsets of a vertex set V with $|V| = n$. If G is a random d-regular graph on V then the expected value edges between S and T is $d|S||T|/n$. For a graph G, one can ask what the difference is between this expected value and $|E(S,T)|$. The following lemma shows that this deviation, or discrepancy as it is sometimes called, can be estimated by $\mu(G)$.

Lemma 2.21 (Expander mixing lemma)**.** *Let G be a d-regular graph on n vertices. Then for all $S, T \subset V$:*

$$\left| |E(S,T)| - \frac{d|S||T|}{n} \right| \leq \mu(G)\sqrt{|S||T|}$$

Proof. Let χ_S and χ_T be the characteristic functions of S and T. Then $|E(S,T)| = \langle A\chi_S, \chi_T \rangle$. Expand these functions in the orthonormal basis of eigenfunctions $\phi_1, \dots \phi_n$ where $\phi_1(v) = \frac{1}{\sqrt{n}}$ for every $v \in V$.

$$\chi_S = \sum_{j=0}^{n-1} s_j \phi_j \text{ and } \chi_T = \sum_{j=0}^{n-1} t_j \phi_j$$

In this case $|S| = \langle \chi_S, \chi_S \rangle = \sum_{j=0}^{n-1} s_j^2$ and similarly $|T| = \sum_{j=0}^{n-1} t_j$. In addition $s_0 = \langle \chi_S, \phi_0 \rangle = \frac{|S|}{\sqrt{n}}$ and $t_0 = \frac{|T|}{\sqrt{n}}$. Finally

$$|E(S,T)| = \langle A\chi_S, \chi_T \rangle = \left\langle A\Big(\sum_{j=0}^{n-1} s_j\phi_j\Big), \Big(\sum_{j=0}^{n-1} t_j\phi_j\Big) \right\rangle = \sum_{j=0}^{n-1} \mu_j s_j t_j$$

Considering the previous identities one can see that the first term in the sum on the right hand side, ds_0t_0, equals $d\frac{|S||T|}{n}$. Hence

$$\left| |E(S,T)| - \frac{d|S||T|}{n} \right| = \left| \sum_{j=1}^{n-1} \mu_j s_j t_j \right| \leq \mu(G) \sum_{j=1}^{n-1} |s_j||t_j|$$

Applying the Cauchy-Schwarz inequality for the sum and taking into account the missing terms $|s_0|^2$ and $|t_0|^2$ we get the required upper bound.

$$\mu(G)\Big(\sum_{1\leq j\leq n-1}|s_j|^2\Big)^{1/2}\Big(\sum_{1\leq j\leq n-1}|t_j|^2\Big)^{1/2}=\mu(G)\sqrt{|S||T|}$$

□

Proposition 2.22 (Upper bound for independence number)**.** *For an* (n,d)*-graph* G

$$\alpha(G)\leq\frac{\mu(G)}{d}n$$

Proof. This statement is an immediate consequence of the expander mixing lemma (Lemma 2.21). The independence property of a set $I\subset V$ is equivalent to the condition $E(I,I)=\emptyset$. The lemma provides the appropriate estimation with the choice $S=T=I$. □

Proposition 2.23 (Lower bound for chromatic number)**.** *For an* (n,d)*-graph* G

$$\chi(G)\geq\frac{d}{\mu(G)}$$

Proof. A k-coloring of a graph $G=(V,E)$ is a function $c\in L^2(G,\mu)$ such that $c(V)=\{1,2,\dots k\}$ and $c(v)\neq c(w)$ for two adjacent vertices v and w. For every $j\in\{1,2,\dots k\}$ the set $c^{-1}(\{j\})$ is an independent set in G. Consequently $n=\sum_{j\leq k}\left|c^{-1}(\{j\})\right|\leq\sum_{j\leq k}\frac{\mu(G)}{d}n\leq\chi(G)\frac{\mu(G)}{d}n$. □

Proposition 2.24 (Upper bound for diameter)**.** *For an* (n,d)*-graph* G

$$\operatorname{diam}(G)\leq\frac{\log(2n)}{\log\left(\frac{d+\sqrt{d^2-\mu(G)^2}}{\mu(G)}\right)}$$

Proof. The proof presented here is can be found in Sarnak's book [Sar90]. Let A be the adjacency operator of G. Furthermore let $\{\phi_j\}_{j=0}^{n-1}$ be an orthonormal basis of $L^2(G,\mu)$ consisting of eigenfunctions of A. As in the previous proofs, $\phi_0(v)=\frac{1}{\sqrt{n}}$ for every $v\in V$. Applying the spectral theorem (Theorem 2.13), we have

$$Af=\sum_{j=0}^{n-1}\mu_j\langle\phi_j,f\rangle\phi_j\text{ and }P(A)f=\sum_{j=0}^{n-1}P(\mu_j)\langle\phi_j,f\rangle\phi_j\text{ for every }f\in L^2(G,\mu).$$

for any polynomial P. Let us introduce χ_v as the characteristic function of the set consisting only the single vertex v. Then

$$\big(P(A)\chi_v\big)(w)=\sum_{j=0}^{n-1}P(\mu_j)\langle\phi_j,\chi_v\rangle\phi_j(w)=\sum_{j=0}^{n-1}P(\mu_j)\phi_j(v)\phi_j(w).$$

If one applies the A^k operator to χ_v function, then $A^k\chi_v$ counts the walks of length k in G starting from v. In particular, its value at a vertex w is the number of walks of appropriate length which start in v and end in w. Hence, if $d(v,w)>N$ and $\deg(P)\leq N$, then clearly $\big(P(A)\chi_v\big)(w)=0$. This yields

$$0=\sum_{j=0}^{n-1}P(\mu_j)\phi_j(v)\phi_j(w)$$

Rearranging the terms and taking absolute value we get

$$\frac{P(d)}{n} = \left|\sum_{j=1}^{n-1} P(\mu_j)\phi_j(v)\phi_j(w)\right| \leq \Big(\sup_{|\mu|\leq\mu(G)} |P(\mu)|\Big) \sum_{j=1}^{n-1} |\psi_j(v)\psi_j(w)|.$$

At this point we can apply the inequality of quadratic and geometric means for the terms in the summation in order to reduce the question to an extremal problem in analysis.

$$\frac{P(d)}{n} \leq \Big(\sup_{|\mu|\leq\mu(G)} |P(\mu)|\Big)\frac{1}{2}\sum_{j=1}^{n-1}\big(|\phi_j(v)|^2 + |\phi_j(w)|^2\big)$$

Combining the previous calculations, we have $P(d) \leq n\sup_{|\mu|\leq\mu(G)}|P(\mu)|$ for any polynomial of degree at most N. We can apply this inequality with $T_N(x/\mu(G))$, where T_N is the N-th Chebyshev polynomial of the first kind, i.e. $T_N(\cos x) = \cos(Nx)$ for $|x| \leq 1$. These polynomials satisfy the following conditions for $-1 \leq \cos t = x \leq 1$ by definition.

$$T_N(\cos t) + i\sin(Nt) = (\cos t + i\sin t)^N = \big(x + i\sqrt{1-x^2}\big)^N$$
$$T_N(\cos t) - i\sin(Nt) = (\cos t - i\sin t)^N = \big(x - i\sqrt{1-x^2}\big)^N$$

Combining these equations we get an explicit formula for the polynomials.

$$T_n = \tfrac{1}{2}\Big(\big(x + i\sqrt{1-x^2}\big)^N + \big(x - i\sqrt{1-x^2}\big)^N\Big)$$

Hence we get from the derived inequality that $T_N\big(\frac{x}{\mu(G)}\big) \leq n$. Since the powers in the explicit formula for Chebyshev polynomials are complex conjugates, giving a real lower bound for one of the terms provides us an appropriate estimate, namely

$$\frac{1}{2}\left(\frac{d + \sqrt{d^2 - \mu(G)^2}}{\mu(G)}\right)^N \leq n.$$

Or equivalently

$$N \leq \frac{\log(2n)}{\log\left(\frac{d+\sqrt{d^2-\mu(G)^2}}{\mu(G)}\right)}.$$

□

2.3 Asymptotic behaviour of eigenvalues

As we have seen in Corollary 2.19 , the quality of an expander family can be measured by a lower bound on the spectral gap. On the other hand, it turns out that, asymptotically, the spectral gap cannot be too large. All the graphs in this section are supposed to be without loops.

2.3.1 Spectrum of the infinite tree

As we have seen, the expansion constant is strongly related to the spectral gap. We start our discussion with the spectral analysis of extremal instances. While the finite d-regular trees have poor expander properties, the infinite d-regular tree $\mathbb{T}_d$ is the perfect expander. The question of estimating expander constant in this viewpoint is how close one can get to this level of expansion with finite d-regular graphs.

Theorem 2.25 (Cartier)**.** *The spectrum of infinite d-regular tree $\mathbb{T}_d$ is*

$$\mathrm{Sp}(A) = [-2\sqrt{d-1}, 2\sqrt{d-1}].$$

Proof. The proof presented here is based on Sunada's article [Sun88]. A similar proof can be found in [Fri91]. Let $\mathbb{T}_d = (V, E)$. The adjacency operator A is self-adjoint, hence by Theorem 2.11 and Proposition 2.10 we have

$$r_A = \lim_{n\to\infty} \|A^n\|^{1/n} = \sup_{\mu\in\mathrm{Sp}(A)} |\mu| = \|A\|.$$

This means that the spectral radius is the following

$$\sup_{\substack{f\in L^2(\mathbb{T}_d,\mu)\\ f\neq 0}} \frac{\left|\int_V (Af)(x)\overline{f(x)}\,\mathrm{d}\mu(x)\right|}{\int_V |f(x)|^2\,\mathrm{d}\mu(x)} = \sup_{\substack{f\in L^2(\mathbb{T}_d,\mu)\\ f\neq 0}} \frac{|\langle Af, f\rangle|}{\langle f, f\rangle} = \|A\| = \sup_{\mu\in\mathrm{Sp}(A)} |\mu|$$

An orientation of an edge $e = (x, y)$ is a bijective function $\varphi_e\{x, y\} \to \{\pm 1\}$. An orientation of edges means that there is an orientation φ_e for every $e \in E$. We employ the following notation: $e_+ = \varphi_e^{-1}(1)$ and $e_- = \varphi_e^{-1}(-1)$. An orientation can be assigned to each edge in such a way that for each vertex $v \in V$, there exists only one oriented edge with $e_- = v$. This assignment is possible because there are no circles in $\mathbb{T}_d$. It is sufficient to prove that $\left|\langle Af, f\rangle\right| \le 2\sqrt{d-1}\|f\|^2$ for all finitely supported function f on V. Then we have

$$|\langle Af, f\rangle| = \Bigg|\sum_{\substack{v\in V\\ w\in N(v)}} f(v)\overline{f(w)}\Bigg| = \Bigg|\sum_{e\in E}\big(f(e_-)\overline{f(e_+)} + f(e_+)\overline{f(e_-)}\big)\Bigg| = 2\Bigg|\sum_{e\in E}\Re\big(f(e_+)\overline{f(e_-)}\big)\Bigg|$$

$$\le 2\sum_{e\in E}\left|f(e_+)\right|\left|f(e_-)\right| \le 2\Big(\sum_{e\in E}\left|f(e_+)\right|^2\Big)^{\frac{1}{2}}\Big(\sum_{e\in E}\left|f(e_-)\right|^2\Big)^{\frac{1}{2}}$$

Here our assumption on the orientation can be used. It implies that the correspondence $e \to e_-$ is one-to-one, while the correspondence $e \to e_+$ is $(d-1)$-to-one.

$$|\langle Af, f\rangle| \le 2\Big(\int_V |f(x)|^2\,\mathrm{d}\mu(x)\Big)^{\frac{1}{2}}(d-1)^{\frac{1}{2}}\Big(\int_V |f(x)|^2\,\mathrm{d}\mu(x)\Big)^{\frac{1}{2}} = 2\sqrt{d-1}\|f\|^2$$

This proves our first claim, namely that the spectral radius is at most $2\sqrt{d-1}$. In order to prove that it is at least this much, for every $\varepsilon > 0$ we have to find a function f such that $\langle Af, f\rangle \ge \big(2\sqrt{d-1}-\varepsilon\big)\|f\|^2$. For a $v_0 \in V$ and $n \in \mathbb{Z}^+$ let $f_{v_0,n}(v) \equiv f(v)$ be defined as

$$f(v) = \begin{cases} 1 & \text{if } v = v_0 \\ \frac{1}{\sqrt{d(d-1)^{k-1}}} & \text{if } d(v, v_0) = k \text{ with } 1 \le k \le n-1 \\ 0 & \text{if } d(v, v_0) \ge n. \end{cases}$$

It can be seen immediately that $\|f\| = n$. Then for $\langle Af, f\rangle$ we have

$$\sum_{v\in V}\sum_{w\in N(v)} f(v)f(w) = 2\sqrt{d} + 2\sum_{k=1}^{n-2} d(d-1)^{k-1}\frac{1}{\sqrt{d(d-1)^{k-1}}}\frac{d-1}{\sqrt{d(d-1)^k}} \ge 2(n-1)\sqrt{d-1}.$$

And therefore

$$\lim_{n\to\infty}\frac{2(n-1)\sqrt{d-1}}{n} = 2\sqrt{d-1} \le \sup_{\substack{f\in L^2(\mathbb{T}_d,\mu)\\ f\neq 0}}\frac{|\langle Af, f\rangle|}{\langle f, f\rangle},$$

which was to be demonstrated. □

Another property of $\mathbb{T}_d$ will be used, namely it is the universal covering space of every finite d-regular graph. We will return to this topic in Chapter 4. At the moment the only important fact for us is that the number of walks in a d-regular graph can be estimated by the same quantity as in $\mathbb{T}_d$.

Lemma 2.26 (Counting closed walks in $\mathbb{T}_d$)**.** *Let t_s be the number of closed walks of length s that start and end at some given vertex r in $\mathbb{T}_d$. Then for every $s \in \mathbb{Z}^+$*

(i) $t_{2s+1} = 0$.

(ii) $t_{2s} \geq 2\sqrt{d-1}\left(1 - O\left(\frac{\log k}{k}\right)\right)$.

(iii) $t_{2s} = \sum_{j=1}^{s} \binom{2s-j}{s} \frac{j}{2s-j} d^j (d-1)^{s-j}$.

Proof. (i) The first claim follows from the fact that a tree is a bipartite graph which means it contains no odd length cycles.

(ii) Every walk that starts and ends at the same vertex r in $\mathbb{T}_d$ can be associated with a sign pattern $\varphi : \{1, 2, \dots 2s\} \to \{-1, 1\}$. Each step of the walk can be associated ± 1 according to whether it is directed either away or towards r and the partial sum of each prefix is non-negative (since it describes the distance between our current position and the root). It is well-known that the number of such functions is the s-th Catalan number $C_s = \frac{1}{s+1}\binom{2s}{s}$. (This can be verified via André's reflection principle for example.) For each function φ with the properties listed above, there are at least $(d-1)^s$ different walks, since there are exactly s occurrences of $+1$ in the sequence when there are at least $d-1$ options for directions to continue the walk without changing the sign pattern. In summary, $t_{2s} \geq \frac{1}{s+1}\binom{2s}{s}(d-1)^s$.

(iii) For the purpose of verifying the exact number of such walks, we only have to fine-tune the previous argument. The number of such sequences which represent a walk with exactly j recurrences is $\frac{j}{2s-j}\binom{2s-j}{s}$. In a step towards r, the next vertex is uniquely determined, while on steps away from r, we have d choices whenever our position is r and $d-1$ choices otherwise. Summing these cases provides us with the required equation. □

2.3.2 Trace formula for graphs

In this section we will follow the treatment of Sarnak in [DSV03]. Let $G = (V, E)$ be a d-regular graph. A path of length k without backtracking in G is a sequence $(v_0, v_1, \dots v_k)$ of vertices such that $v_{j+1} \in N(v_j)$ if $j \leq k-1$ and $v_{j+1} \neq v_{j-1}$ if $1 \leq j \leq k-1$. For $v, w \in V$ denote by $p_k(v, w)$ the number of paths of length k without backtracking with origin v and endpoint w. If $v = w$ then we abbreviate this quantity by $p_k(v)$. In addition, for every $k \in \mathbb{N}$ we introduce the operator $A_k : L^2(G, \mu) \to L^2(G, \mu)$ defined by the condition $\langle A_k \chi_v, \chi_w \rangle = p_k(v, w)$. If one represents these operators with matrices indexed by $V \times V$ then the condition simply means that the corresponding matrix entries are $p_k(v, w)$. Note that $A_0 = I$ and $A_1 = A$, the adjacency operator of G.

Proposition 2.27. *Let $\{A_k\}_{k \in \mathbb{N}}$ be the operators defined above and A the adjacency operator of the d-regular graph G. Then*

(i) $A_1^2 = A_2 + dI$.
(ii) $A_1 A_k = A_{k+1} + (k-1)A_{k-1}$ *for* $k \geq 2$.
(iii) $A_1 A_k = A_k A_1$.
(iv) *in the ring* $\operatorname{End}\big(L^2(G,\mu)\big)[[x]]$ *of formal power series over* $\operatorname{End}\big(L^2(G,\mu)\big)$ *the following identity holds* $\Big(\sum_{0\leq k} A_k x^k\Big)\big(I - Ax + (d-1)x^2 I\big) = (1-d^2)I$

Proof. Since we would like to take advantage of the combinatoric side of the definitions, it will be more convenient to work with matrices rather than operators. In the light of this, let us assume that A_k and A are matrices indexed by $V \times V$.

(i) For $v, w \in V$ the entry $\big(A^2\big)_{vw}$ is the number of all walks of length 2 between v and w. The difference between these quantities and values $p_k(v,w)$ occurs when $v = w$ because when $v \neq w$ then backtracking is not even possible during a walk. If $v = w$, then $p_k(v,w)$ does not involve the walks taken from v to one of its neighbours and back.

(ii) $\big(A_k A_1\big)_{vw}$ is the number of paths of length $k+1$ from v to w without backtracking except possibly in the last step. Hence the set of such walks can be partitioned into two classes according to the direction of the last step. There are $\big(A_{k+1}\big)_{vw}$ walks without any backtracking. If there is a backtracking at the last step, w has to be reached in $k-1$ steps which can be done in $(A_{k-1})_{vw}$ different ways and there are still $d-1$ possibilities to step one forward and back to finish the walk.

(iii) In the light of the previous claim and argument this can be seen by changing the direction of every walk.

(iv) This is a straightforward calculation based on the previous claims of the proposition. □

For the purpose of revealing the generating function of the operator series $\{A_k\}_{0\leq k}$, the last identity can be reformulate as $\sum_{0\leq k} A_k x^k = \frac{1-x^2}{1-Ax+(d-1)x^2}$. Let us introduce the operators T_m for $m \in \mathbb{N}$ defined by $\sum_{0\leq k\leq \frac{m}{2}} A_{m-2k}$. The generating function of these operators will be more convenient because of its trivial numerator.

$$\sum_{0\leq m} T_m x^m = \sum_{0\leq k} x^{2k} \sum_{2k\leq m} A_{m-2k} x^{m-2k} = \left(\frac{1}{1-x^2}\right)\left(\frac{1-x^2}{1-Ax+(d-1)x^2}\right).$$

Proposition 2.28. *For* $m \in \mathbb{N}$ *let* U_m *denote the* m*-th Chebyshev polynomial of the second kind. Then*

$$T_m = (d-1)^{\frac{m}{2}} U_m\left(\frac{A}{2\sqrt{d-1}}\right)$$

Proof. By definition $U_m(\cos t) = \frac{\sin(m+1)t}{\sin t}$ hence the polynomials satisfy the recurrence relation $U_{m+1}(z) = 2xU_m(z) - U_{m-1}(z)$. Applying this recurrence relation the generating function of the polynomials can be computed easily

$$\big(1-2zx+x^2\big)\sum_{m=0}^{\infty} U_m(z)x^m = U_0(z) + \big(U_1(z)-2z\big) + \sum_{m=2}^{\infty}\big(U_{m+1}(z) - 2zU_m(z) + U_{m-1}(z)\big)x^m = 1.$$

An explicit formula for the generating function results from rearranging the sides of the above equation.

$$\sum_{m=0}^{\infty} U_m(z)x^m = \frac{1}{1-2zx+x^2}.$$

Performing a change of variables with the substitution $x := (d-1)^{1/2}x$ and $z := \frac{A}{2\sqrt{d-1}}$ we have

$$\sum_{m=0}^{\infty}(d-1)^{\frac{m}{2}}U_m\left(\frac{A}{2\sqrt{d-1}}\right)x^m = \frac{1}{1-A+(d-1)x^2} = \sum_{m=0}^{\infty}T_m x^m$$

□

Theorem 2.29 (Trace formula for graphs)**.** *Let $p_k(x)$ be the number of paths of length k without backtracking from x to x and let U_m denote the m-th Chebyshev polynomial of the second kind. Then*

$$\int_V \Big(\sum_{0\leq k\leq \frac{m}{2}} p_{m-2k}(x)\Big)\, \mathrm{d}\mu(x) = (d-1)^{\frac{m}{2}}\sum_{j=0}^{n-1}U_m\left(\frac{\mu_j}{2\sqrt{d-1}}\right)$$

where μ_j is an eigenvalue of the adjacency operator A.

Proof. Although the operator T_m is not present in the statement, the proposition concerns its trace, hence calculating this quantity in two different ways will lead us to the claimed trace formula for graphs. As a consequence of the spectral theorem (Theorem 2.13), $\mathrm{Tr}(T_m)$ is the sum of the eigenvalues of T_m. Since the previous lemma expresses T_m as a polynomial of degree m of the adjacency operator A, the trace can be calculated as

$$\mathrm{Tr}(T_m) = (d-1)^{\frac{m}{2}}\sum_{j=0}^{n-1}U_m\left(\frac{\mu_j}{2\sqrt{d-1}}\right).$$

Otherwise, by the definition of T_m we have

$$\mathrm{Tr}(T_m) = \sum_{0\leq k\leq \frac{m}{2}} \mathrm{Tr}(A_{m-2k}) = \sum_{0\leq k\leq \frac{m}{2}} \int_V p_{m-2k}(x)\, \mathrm{d}\mu(x)$$

□

2.3.3 Asymptotic formula

Now we are ready to establish asymptotic results for eigenvalues. It was already proven in Proposition 2.20 that $\mu(G) \geq \sqrt{d}\big(1-o(1)\big)$. Here we prove a stronger bound.

Theorem 2.30 (Alon-Boppana)**.** *Let G be an (n, d)-graph. Then*

$$\mu(G) \geq 2\sqrt{d-1} - o_n(1)$$

The proof illustrated here ([HLW06]) achieves a little less than the original theorem claims because the resulting error term will be slightly weaker. On the other hand, the argument is a straightforward continuation of the previously demonstrated ideas.

Proof. As noted, if $Af = \mu f$ then $A^{2k}f = \mu^{2k}f$, which implies that the maximum of absolute values of eigenvalues of A^{2k} except d^{2k} is exactly $\mu(G)^{2k}$.

Let $v, w \in V$ vertices at distance $\mathrm{diam}(G)$ and let us define the function $\psi = \chi_v - \chi_w$ where χ_v and χ_w are the characteristic functions of sets containing the corresponding vertices. In this case, $\langle \psi, \phi_0 \rangle = 0$, i.e. ψ is orthogonal to the constant function. We have

$$\max_{1\le j\le n-1} \mu_j^{2k} \ge \frac{\langle A^{2k}\psi, \psi\rangle}{\langle \psi, \psi\rangle} = \tfrac{1}{2}\Big(\langle A^{2k}\chi_v, \chi_v\rangle - \langle A^{2k}\chi_v, \chi_w\rangle + \langle A^{2k}\chi_w, \chi_w\rangle - \langle A^{2k}\chi_w, \chi_v\rangle\Big).$$

If one chooses k to be $\lfloor \mathrm{diam}(G)/2\rfloor - 1$, then the negative terms will disappear since there will be no walk of length $2k$ between the two vertices which have distance $d(v, w) = \mathrm{diam}(G)$. On the other hand, for the number of closed walks counted by the positive terms the analogous quantity in $\mathbb{T}_d$, namely t_{2k}, is a suitable lower bound. In summary, $\mu(G)^{2k} \ge t_{2k}$. By Lemma 2.26, we have

$$\mu(G)^{2k} \ge \frac{1}{k+1}\binom{2k}{k}(d-1)^k = (d-1)^k \frac{2^{2k}}{\sqrt{\pi}k^{3/2}}\Big(1 + O(1/k)\Big)$$

$$\mu(G) \ge 2\sqrt{d-1}\Big(1 - O\big(\tfrac{\log k}{k}\big)\Big) = 2\sqrt{d-1}\Big(1 - O\big(\tfrac{\log \mathrm{diam}(G)}{\mathrm{diam}(G)}\big)\Big)$$

Since there exists $c > 0$ such that for every sufficiently large n $\mathrm{diam}(G) \ge c \log n$ holds, it follows that

$$\mu(G) \ge 2\sqrt{d-1}\Big(1 - O\big(\tfrac{\log\log n}{\log n}\big)\Big) = 2\sqrt{d-1} - o_n(1)$$

□

As we have seen, the fact that the number of paths of length k from a vertex v to v in a d-regular graph is at least the number of such walks started and ended in the same vertex in the infinite d-regular tree $\mathbb{T}_d$ has important consequences. In the following sections, we will elaborate on this relation to provide a stronger result. The trace formula for graphs (Theorem 2.29) will be our main tool for this purpose. The improvement of the previous result claims that not only the second largest eigenvalue (in absolute value) becomes asymptotically larger than $2\sqrt{d-1}$, but also a positive proportion of eigenvalues lies in any interval $[(2-\varepsilon)\sqrt{d-1}, d]$. In order to conclude this theorem, two additional technical lemmas will be necessary which highlight us an interesting connection between Chebyshev polynomials and certain probability measures. It will be more convenient to introduce the notation u_m for polynomials $U_m\big(\frac{x}{2}\big)$. These polynomials u_m satisfy the simpler recursion formula

$$u_{m+1}(x) = x u_m(x) - u_{m-1}(x).$$

Clearly by definition, $u_m(2\cos t) = \frac{\sin(m+1)t}{\sin t}$ and $u_m\big(2\cos\frac{k\pi}{m+1}\big) = 0$ for $1 \le k \le m$. In particular, the largest root of $u_m(x)$ is $\alpha_m = 2\cos\big(\frac{\pi}{m+1}\big)$.

Lemma 2.31. *Let $\varepsilon > 0$ and $L \ge 2$ be fixed. For every probability measure ν on $[-L, L]$ such that $\int_{-L}^{L} u_m(x) d\nu(x) \ge 0$ holds for every $m \in \mathbb{N}$, we have $\nu\big([2-\varepsilon, L]\big) > 0$.*

Proof. Assume by contradiction that $\nu\big([2-\varepsilon, L]\big) = 0$. Take m large enough to have $2-\varepsilon < \alpha_m$. This can be done since the sequence α_m increases to 2. Set $g_m(x) = \frac{u_m(x)^2}{x-\alpha_m}$. If $x \le 2-\varepsilon < \alpha_m$, then $g_m(x) \le 0$ hence

$$\int_{-L}^{L} g_m(x)\,\mathrm{d}\nu(x) = \int_{-L}^{2-\varepsilon} g_m(x)\,\mathrm{d}\nu(x) \le 0$$

In fact, this integral is supposed to be zero. In order to prove this, we need the fact that $g_m(x)$ can be expressed as a linear combination of the polynomials $u_m(x)$ with non-negative coefficients, i.e. $g_m(x) = \sum_{j=1}^{2m-1} c_j u_j(x)$ where $c_j \ge 0$. First we prove that $g_m(x)$ can be written as $\sum_{j=0}^{m-1} u_{m-1-j}(\alpha_m)u_j(x)u_m(x)$. In this form, the coefficients are non-negative, since if $j < m$ then $\alpha_m > \alpha_j$ and $u_j(\alpha_m) > 0$ because α_j was the largest root of u_j. It is sufficient to prove that

$$\left(\sum_{j=0}^{m-1} u_{m-1-j}(\alpha_m)u_j(x)u_m(x)\right)(x-\alpha_m) = u_m(x)$$

This can be shown by a straightforward calculation using the recursion formula for the polynomials $u_m(x)$. Another recurrence relation, namely

$$\sum_{k=0}^{j} u_{m+j-2k} = u_m(x)u_j(x) \text{ for } j \leq m$$

can be proved by induction on j. Combining the formulas we have

$$g_m(x) = \sum_{j=0}^{m-1} u_{m-1-j}(\alpha_m)u_j(x)u_m(x) = \sum_{j=0}^{m-1} u_{m-1-j}(\alpha_m)\sum_{k=0}^{j} u_{m+j-2k}(x)$$

$$g_m(x) = \sum_{j=0}^{2m-1} P_j\big(u_0(\alpha_m), u_1(\alpha_m)\dots u_{m-1}(\alpha_m)\big)u_j(x)$$

where $P_j \in \mathbb{Z}[x_0, x_1, \dots x_{m-1}]$ with non-negative coefficients. As we noted, $u_l(\alpha_m) > 0$ for $l < m$, hence the corresponding polynomial expressions in the last summation are non-negative which was to be proven, i.e. $g_m(x) = \sum_{j=1}^{2m-1} c_j u_j(x)$ where $c_j \geq 0$. By this expression, we have

$$\int_{-L}^{L} g_m(x)\,\mathrm{d}\nu(x) \geq \sum_{j=1}^{2m-1} c_j \int_{-L}^{L} u_m(x)\,\mathrm{d}\nu(x) \geq 0.$$

In comparison to the lower bound for the integral this implies that

$$\operatorname{supp}\nu \subset \bigcap_{m=1}^{\infty} \big\{x \in [-L, L] \mid g_m(x) = 0\big\} = \emptyset.$$

which results a contradiction. □

Lemma 2.32. *Let $\varepsilon > 0$ and $L \geq 2$ be fixed. There exists a constant $C(\varepsilon, L) > 0$ with the following property: for any probability measure ν on $[-L, L]$, such that $\int_{[-L,L]} u_m(x)d\nu(x) \geq 0$ for every $m \in \mathbb{N}$, we have $\nu\big([2-\varepsilon, L]\big) \geq C(\varepsilon, L)$.*

Proof. The compact normed space $\big([-L, L], |.|\big)$ will be denoted by X. Let $\psi \in \mathcal{C}(X)$ be a continuous function with the following properties: $\operatorname{supp}\psi = [2-\varepsilon, L+\varepsilon]$ and if $x \in [2-\varepsilon/2, L]$ then $\psi(x) = 1$. This function will be used to continuously approximate the characteristic function of the set $[2-\varepsilon, L]$. Let us introduce the notation $\mathcal{P}_X$ for the set of probability measures on X which satisfy the condition formulated in the statement, i.e.

$$\mathcal{P}_X = \left\{\nu : \mathcal{B}(X) \to [0,1] \;\middle|\; \nu \text{ probability measure on } X \text{ and } \int_X u_m(x)\,\mathrm{d}\nu(x) \geq 0\ \forall m \in \mathbb{N}\right\}.$$

For every $\nu \in \mathcal{P}_X$, considering the fixed properties of ψ, we have

$$\nu\big([2-\varepsilon, L]\big) = \int_{[2-\varepsilon,L]} 1\,\mathrm{d}\nu(x) \geq \int_X \psi(x)\,\mathrm{d}\nu(x) \geq \int_{[2-\varepsilon/2,L]} 1\,\mathrm{d}\nu(x) \geq \nu\big([2-\varepsilon/2, L]\big)$$

which is larger than an suitable $C_\nu > 0$ by the previous lemma. Let $\mathcal{M}(X)$ denote the set of all probability measures on the measurable space $\big(X, \mathcal{B}(X)\big)$, equipped with the weak topology. As a consequence of the Riesz Representation Theorem, the compactness of X implies the compactness of $\mathcal{M}(X)$. Furthermore, because $\mathcal{P}_X$ is closed subset of $\mathcal{M}(X)$ hence $\mathcal{P}_X$ is also compact in the weak topology. Since ψ is continuous, the map $\mu_\psi : \mathcal{P}_X \to \mathbb{R}^+$ defined by

$$\mu_\psi : \nu \to \int_X \psi \, \mathrm{d}\nu(x)$$

is weakly continuous. By the compactness of $\mathcal{P}_X$, from the open cover

$$\bigcup_{\nu \in \mathcal{P}} \mu_\psi^{-1}\big((C_\nu, 1+\varepsilon)\big)$$

one can choose a finite subcover with corresponding constants $C_1, C_2, \ldots C_l$. In summary, if $C(\varepsilon, L) = \min\{C_1, \ldots, C_l\}$, then $\nu\big([2-\varepsilon, L]\big) \geq C(\varepsilon, L)$ for every $\nu \in \mathcal{P}_X$ which was to be demonstrated. □

Theorem 2.33 (Serre)**.** *For every $\varepsilon > 0$, there exists a constant $C(\varepsilon, d) > 0$ such that, for every finite, connected (n,d)-graph G, the number of eigenvalues of G in the interval $[(2-\varepsilon)\sqrt{d-1}, d]$ is at least $C(\varepsilon, d)n$.*

Proof. Let δ_s be the Dirac measure concentrated on $\{s\}$, i.e. $\displaystyle\int_X f(x)\mathrm{d}\delta_s(x) = f(s)$, for every $f \in \mathcal{C}(X)$. Take

$$L = \frac{d}{\sqrt{d-1}} \text{ and } \nu = \frac{1}{n}\sum_{j=0}^{n-1} \delta_{\left(\frac{\mu_j}{\sqrt{d-1}}\right)}.$$

By these choices, ν is a probability measure on $X = \big([-L, L], |.|\big)$. By the trace formula

$$\int_X u_m(x)\, \mathrm{d}\nu(x) = \frac{1}{n}\sum_{j=0}^{n-1} U_m\left(\frac{\mu_j}{2\sqrt{d-1}}\right) = \int_V \Big(\sum_{0 \leq k \leq \frac{m}{2}} p_{m-2k}(x)\Big)\, \mathrm{d}\mu(x) \geq 0.$$

So $\nu \in \mathcal{P}_X$ and hence the previous lemma is applicable. Therefore there exists $C(\varepsilon, d) > 0$ such that $\nu\big([2-\varepsilon, L]\big) \geq C(\varepsilon, d)$. On the other hand

$$\nu\big([2-\varepsilon, L]\big) = \int_X \chi_{[2-\varepsilon, L]}(x)\, \mathrm{d}\nu(x) = \frac{1}{n}\sum_{j=0}^{n-1}\int_X \chi_{[2-\varepsilon, L]}(x)\mathrm{d}\delta_{\left(\frac{\mu_j}{\sqrt{d-1}}\right)}.$$

$$C(\varepsilon, d)n \leq \left|\left\{j \mid 0 \leq j \leq n-1, (2-\varepsilon) \leq \tfrac{\mu_j}{\sqrt{d-1}} \leq L\right\}\right|$$

And the right-hand side of the last equation is nothing else but the number of eigenvalues of A in the interval $\big[(2-\varepsilon)\sqrt{d-1}, d\big]$, which was to be demonstrated. □

Let $\mu_{n-1}(G)$ denote the smallest eigenvalue of graph G. By a little modification of the above argument one can show that if $(G_n)_{n\in\mathbb{Z}^+}$ is a family of connected, d-regular, finite graphs with $\lim_{n\to\infty} \mathrm{girth}(G_n) = \infty$, then $\limsup_{m\to\infty} \mu_{n-1}(G) \geq -2\sqrt{d-1}$. For proof, see [DSV03].

3. Automorphic forms

There are five elementary arithmetical operations: addition, subtraction, multiplication, division, and... modular forms.

attributed to Martin Eichler

3.1 Preliminaries

Automorphic forms are not only central objects of number theory, but belong to those few mathematical concepts that possess the wonderful ability to link the diverging branches of mathematics. Amazing achievements of mathematics were born by successfully relating certain mathematical objects to automorphic forms, or more specifically to modular forms. For example the Modularity Theorem, or as formerly called the Shimura-Taniyama-Weil conjecture, states that all rational elliptic curves arise from modular forms. The theorem was proved for semistable elliptic curves by Wiles, completing the proof of Fermat's Last Theorem. Despite the beauty of the general theory of automorphic forms, only the concept of modular forms will be required for our purposes but there is still some preparation needed for this.

3.1.1 Riemann surfaces

Definition 3.1 (Riemann surface)**.** A Riemann surface is a one-complex dimensional connected complex analytic manifold, that is, a two-real-dimensional connected manifold $\mathcal{R}$ with a maximal set of charts $\{\varphi_i\}_{i\in I}$ and a collection of sets $\{U_i\}_{i\in I}$ constituting an open cover of $\mathcal{R}$ where $\varphi_i : U_i \to \mathbb{C}$ is homeomorphism onto an open subset of the complex plane $\mathbb{C}$ such that the transition functions

$$\varphi_{ij} = \varphi_i \circ \varphi_j^{-1} : \varphi_j(U_i \cap U_j) \to \varphi_i(U_i \cap U_j)$$

are holomorphic whenever $U_i \cap U_j \neq \emptyset$.

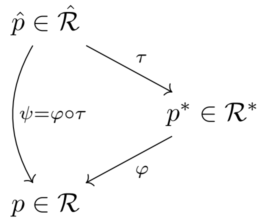
The simplest Riemann surface is the complex plane $\mathbb{C}$. A continuous map $f : \mathcal{R} \to \mathcal{S}$ between Riemann surfaces is called holomorpic (or analytic) if for every local coordinate $\varphi : U \to \mathbb{C}$ on $\mathcal{R}$ and every local coordinate $\psi : V \to \mathbb{C}$ with $U \cap f^{-1}(V) \neq \emptyset$, the map $\psi \circ f \circ \varphi^{-1}$ is holomorphic as a map from a subset of $\mathbb{C}$ to $\mathbb{C}$. A holomorphic map into $\mathbb{C}$ is called a holomorphic function, while a holomorphic map into the Riemann sphere, $\mathbb{C} \cup \{\infty\}$ is called a meromorphic function.

Definition 3.2 (Covering of a Riemann surface)**.** A covering of a Riemann surface $\mathcal{R}$ is a pair $(\mathcal{X}, \varphi)$ where $\mathcal{X}$ is a Riemann surface and $\varphi : \mathcal{X} \to \mathcal{R}$ is a holomorphic map which is a local homeomorphism with the following property: for every $r \in \mathcal{R}$ there exists an open neighbourhood U whose pre-image $\varphi^{-1}(U)$ is a disjoint union of open sets in $\mathcal{X}$ each one of which is homeomorphic to U by φ.

Definition 3.3 (Universal covering of a Riemann surface)**.** A covering $(\mathcal{X}, \varphi)$ is called universal covering if $\mathcal{X}$ is simply connected as a topological space, that is, the fundamental group of $\mathcal{X}$ is trivial, i.e., $\pi_1(\mathcal{X}) = \{1\}$. In this case, $\mathcal{X}$ is called the universal covering space.

Theorem 3.4 (Existence of universal covering)**.** *For every Riemann surface $\mathcal{R}$, there exists a universal covering space $\hat{\mathcal{R}}$.*

An important property of coverings is the fact that the curves on the covered space can be lifted up to the covering space, i.e., every closed curve which is not homotopically trivial on $\mathcal{R}$ lifts to an open curve on $\hat{\mathcal{R}}$, and the curve on $\hat{\mathcal{R}}$ is uniquely determined by the curve on $\mathcal{R}$ and the the point lying over its initial point. As a consequence of this property and the Monodromy Theorem, if $\varphi : \mathcal{R}^* \to \mathcal{R}$ is an arbitrary covering with $\hat{p} \in \hat{\mathcal{R}}$, $p^* \in \mathcal{R}^*$ such that $\psi(\hat{p}) = \varphi(p^*)$, then there exists a unique $\tau : \hat{\mathcal{R}} \to \mathcal{R}^*$ covering with the properties $\psi = \varphi \circ \tau$ and $\tau(\hat{p}) = p^*$.



In this case, $\pi_1(\mathcal{R}^*)$ is isomorphic to a subgroup of $\pi_1(\mathcal{R})$ and hence the covering manifolds of $\mathcal{R}$ are in a bijective correspondence with conjugacy classes of subgroups of $\pi_1(\mathcal{R})$. In this setting, $\hat{\mathcal{R}}$ corresponds to the trivial subgroup of $\pi_1(\mathcal{R})$. Furthermore, if in the previous setting we choose $\mathcal{R}^*$ to be $\hat{\mathcal{R}}$ and $\psi : \hat{\mathcal{R}} \to \mathcal{R}$ is universal covering with $\hat{p}_1, \hat{p}_2 \in \hat{\mathcal{R}}$ and $\psi(\hat{p}_0) = \psi(\hat{p}_1)$, then there exists a unique $\vartheta : \hat{\mathcal{R}} \to \hat{\mathcal{R}}$ conformal map such that $\psi = \psi \circ \vartheta$ and $\vartheta(\hat{p}_0) = \hat{p}_1$. These transformations are called covering transformations or deck transformations and they form a group isomorphic to $\pi_1(\mathcal{R})$.

Definition 3.5 (Free and properly discontinuous action)**.** Let Γ be a group which acts on a Riemann surface $\mathcal{R}$ by conformal automorphisms. The action is free and properly discontinuous if

(i) for all $p \in \mathcal{R}$ there exists a neighbourhood U of p such that for all $p_1, p_2 \in U$, if $\gamma p_1 = p_2$ then γ is the identity and $p_1 = p_2$.
(ii) If two points, $z, w \in \mathcal{R}$ are not Γ-equivalent, i.e., are not in the same orbit, then there are neighbourhoods U and V of z, w respectively such that $\gamma U \cap V = \emptyset$ for all $\gamma \in \Gamma$.

A free and properly discontinuous action is necessarily free, which means that every group element different from the identity is fixed-point free. It is not too difficult to see that the deck transformation group of the universal covering acts freely and properly discontinuously on a Riemann surface.

Proposition 3.6. *Let Γ be a group which acts freely and properly discontinuously on a simple connected Riemann surface $\hat{\mathcal{R}}$ then $\mathcal{R} = \Gamma \backslash \hat{\mathcal{R}}$ is a Riemann surface with universal covering space $\hat{\mathcal{R}}$ and with $\pi_1(\Gamma \backslash \hat{\mathcal{R}}) \cong \Gamma$.*

This proposition shows that a Riemann surface can be recovered from its universal covering space. The surface is just the quotient of the universal covering space under the deck transformation group which can be identified with the fundamental group. If our desire is the classification of the Riemann surfaces (up to conformal equivalence) then doing this via their universal covering spaces seems very natural. Fortunately, it turns out that there are exactly three conformally distinct simply connected Riemann surfaces. One of these is compact, it is conformally equivalent to the Riemann sphere $\mathbb{C} \cup \{\infty\}$ (elliptic model). The non-compact simply connected Riemann surfaces are conformally equivalent to either the entire plane $\mathbb{C}$ (parabolic model) or the upper half plane $\mathbb{H}$ (hyperbolic model). This result is usually referred to as the Uniformization Theorem. The designation of these Riemann surfaces as elliptic, parabolic and hyperbolic comes from Riemannian geometry, where it is natural to endow each of these surfaces with a constant curvature Riemannian metric which is positive, zero or negative respectively.

Theorem 3.7 (Uniformization of arbitrary Riemann surfaces)**.** *Every Riemann surface $\mathcal{R}$ is conformally equivalent to $\Gamma\backslash\hat{\mathcal{R}}$ with $\hat{\mathcal{R}} = \mathbb{C} \cup \{\infty\}$, $\mathbb{C}$ or $\mathbb{H}$ and Γ is a group acting freely and properly discontinuously on $\hat{\mathcal{R}}$. Furthermore, $\Gamma \cong \pi_1(\mathcal{R})$.*

It seems that to exploit the power of uniformization theorem one has to investigate the automorphism groups of simply connected Riemann surfaces. Given any complex numbers a, b, c, d with $ad - bc \neq 0$, we can define the Möbius transformation (or fractional linear transformation) on the Riemann sphere $\mathbb{C} \cup \{\infty\}$ by $z \to \frac{az+b}{cz+d}$ for $z \neq \infty, -d/c$, with the convention that $-d/c$ is mapped to ∞ and ∞ is mapped to a/c. The set of fractional linear transformations is closed under composition. The Möbius transformations can be interpreted as projective linear transformations. Every element of the general linear group $GL_2(\mathbb{C})$ generates a Möbius transformation as the next proposition demonstrates:

Proposition 3.8. *The map $F : GL_2(\mathbb{C}) \to \mathrm{Aut}(\mathbb{C} \cup \{\infty\})$ defined by*

$$F : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \to \left(z \to \frac{az+b}{cz+d}\right)$$

is a surjective homomorphisms with kernel

$$\operatorname{Ker} F = \left\{ \begin{pmatrix} \lambda & \\ & \lambda \end{pmatrix} \;\middle|\; \lambda \in \mathbb{C} \setminus \{0\} \right\}.$$

So two different elements of $GL_2(\mathbb{C})$ generate the same Möbius transformation if and only if they are scalar multiplies. Dividing by a scalar, we can represent a transformation by a matrix of determinant 1. If the corresponding matrix lies in $SL_2(\mathbb{C}) = \{A \in GL_2(\mathbb{C}) \mid \det(A) = 1\}$ then the ambiguity implied by the trivial acting I and $-I$ still will be left.

Theorem 3.9 (Automorphisms of the Riemann sphere)**.** *Let $T : \mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\}$ be a complex diffeomorphism. Then T is a Möbius transformation. Furthermore,*

$$\mathrm{Aut}\left(\mathbb{C} \cup \{\infty\}\right) \cong SL_2(\mathbb{C})/\{I, -I\} \cong PSL_2(\mathbb{C})$$

Theorem 3.10 (Automorphisms of the complex plane)**.** *Let $T : \mathbb{C} \to \mathbb{C}$ be a complex diffeomorphism. Then T is an affine transformation. Furthermore,*

$$\mathrm{Aut}(\mathbb{C}) = \left\{ az + b \mid a \in \mathbb{C} \setminus \{0\}\ b \in \mathbb{C} \right\}$$

Theorem 3.11 (Automorphisms of the upper half-plane)**.** *Let $T : \mathbb{H} \to \mathbb{H}$ be a complex diffeomorphism from the upper half-plane to itself. Then there exists real numbers a, b, c, d with $ad - bc = 1$ such that $T(z) = \frac{az+b}{cz+d}$ for $z \in \mathbb{H}$. Furthermore,*

$$\mathrm{Aut}(\mathbb{H}) \cong SL_2(\mathbb{R})/\{I, -I\} \cong PSL_2(\mathbb{R}).$$

As mentioned before, every connected Riemann surface is the quotient of one of the three model surfaces $\mathbb{C} \cup \{\infty\}, \mathbb{C}$ and $\mathbb{H}$ by a group of complex automorphisms that act freely and properly discontinuously. Depending on which surface is used, these are called surfaces of elliptic type, parabolic type and hyperbolic type respectively.

Elliptic types: The automorphism of $\mathbb{C} \cup \{\infty\}$ are the Möbius transformations. From the fundamental theorem of algebra we see that every Möbius transformation has at least one fixed point. Thus the only group of conformal automorphisms that acts freely on $\mathbb{C} \cup \{\infty\}$ is the trivial group. So the only Riemann surface of elliptic type are those that are complex diffeomorphic to the Riemann sphere.

Parabolic types: The affine transformations $z \to az + b$ have fixed points in $\mathbb{C}$ if $a \neq 1$ so in order to obtain a free action Γ has to be restricted to the transformations $z \to z + b$. Thus we can view Γ as an additive subgroup of $\mathbb{C}$.

Proposition 3.12 (Discrete subgroups of $\mathbb{C}$)**.** *Let Γ be a discrete additive subgroup of $\mathbb{C}$. Then Γ takes on one of the following three forms:*

(i) *(Rank zero case) the trivial group $\{0\}$;*

(ii) *(Rank one case) a cyclic group $\omega\mathbb{Z} = \{n\omega \mid n \in \mathbb{Z}\}$ for some $\omega \in \mathbb{C} \setminus \{0\}$;*

(iii) *(Rank two case) a group $\omega_1\mathbb{Z} + \omega_2\mathbb{Z} = \{n_1\omega_1 + n_2\omega_2 \mid n_1, n_2 \in \mathbb{Z}\}$ for some $\omega_1, \omega_2 \in \mathbb{C} \setminus \{0\}$ with ω_2/ω_1 imaginary.*

From the result above we can conclude that every Riemann surface of parabolic type is complex diffeomorphic to a plane $\mathbb{C}$, a cylinder $\omega\mathbb{Z}\backslash\mathbb{C}$ for some $\omega \in \mathbb{C}\backslash\{0\}$, or a torus $(\omega_1\mathbb{Z}+\omega_2\mathbb{Z})\backslash\mathbb{C}$. It can be seen easily that all cylinders are complex diffeomorphic to the punctured plane $\mathbb{C} \setminus \{0\}$ which can be used as a model for all Riemann surface cylinders. As for the tori, it is an important fact in algebraic geometry and number theory that these objects can be modelled by elliptic curves over $\mathbb{C}$.

Hyperbolic types: A Riemann surface of hyperbolic type is isomorphic to a quotient $\Gamma\backslash\mathbb{H}$ of $\mathbb{H}$ by some group $\Gamma \leq PSL_2(\mathbb{R})$ that acts freely and properly discontinuously. The hyperbolic case is of primary interest and it turns out there is a very large number of possible subgroups of $PSL_2(\mathbb{R})$ obeying the mentioned conditions. Investigating these groups is a fundamental topic in hyperbolic geometry which will be within our scope in the next section.

For the proofs of the claims made in this section or for further information one can consult [FK92] or [Hal12].

3.1.2 Hyperbolic geometry

$GL_2^+(\mathbb{R})$, the group of matrices of positive determinant, acts on the upper half-plane $\mathbb{H}$ by the fractional linear transformations. Dividing by a scalar, we can represent the Möbius transformations by a matrix of determinant 1. Let $\gamma = \left(\begin{smallmatrix} * & * \\ c & d \end{smallmatrix}\right)$ be an arbitrary matrix in $\Gamma := SL_2(\mathbb{R})$ and $z \in \mathbb{H}$. Let us introduce the function $j : \Gamma \times \mathbb{H} \to \mathbb{C}$ defined by $j_\gamma(z) = cz + d$. It can be easily verified that this function satisfies the chain rule, so for any $\gamma, \delta \in \Gamma$ and $z \in \mathbb{H}$ the identity $j_{\gamma\delta}(z) = j_\gamma(\delta z) j_\delta(z)$. We also have $\gamma z - \gamma w = (z-w) j_\gamma(z)^{-1} j_\gamma(w)^{-1}$ and hence $\frac{\partial}{\partial z}\gamma z = j_\gamma(z)^{-2}$. The quantity $|j_\gamma(z)|^{-2} = |cz+d|^{-2}$ is called the deformation factor of γ at z. From this, one can observe that if both z, w are on the curve $\mathcal{C}_\gamma = \{\zeta \in \mathbb{H} \mid |j_\gamma(\zeta) = 1|\}$ then $|\gamma z - \gamma w| = |z - w|$. The set of all points of deformation smaller than 1 can be defined as the exterior of $\mathcal{C}_\gamma$. $\mathcal{C}_\gamma$ is called the isometric circle of γ. Since $|j_\gamma(z)|^2 \Im(\gamma z) = \Im z$, there are three invariant subspaces of $\mathbb{C} \cup \{\infty\}$ with respect to the action of group $SL_2(\mathbb{R})$: the upper half-plane $\mathbb{H}$, the lower half-plane $\overline{\mathbb{H}}$ and the real line $\mathbb{R} \cap \{\infty\}$.

Let T be a conformal diffeomorphism of $\mathbb{H}$ and also let us introduce the notation $T(z) = w$. z_0 is a fixed point of the upper half-plane. The transform $\phi_\zeta : \mathbb{H} \to D(0,1)$ defined by $\phi_\zeta(z) = \frac{z-\zeta}{z-\overline{\zeta}}$ is a complex diffeomorphism from the upper half-plane to the unit disk with $\phi_\zeta(\zeta) = 0$. Since $\phi_{z_0}(z_0) = \phi_{w_0} \circ T(z_0)$, the following equation holds

$$\frac{T(z) - T(z_0)}{T(z) - \overline{T(z_0)}} = \varrho \frac{z - z_0}{z - \overline{z_0}}$$

with a constant ϱ satisfying $|\varrho| = 1$. This implies that

$$\lim_{z \to z_0} \left| \frac{T(z) - T(z_0)}{T(z) - \overline{T(z_0)}} \right| = \frac{|T'(z_0)|}{2\Im(T(z_0))} = \frac{1}{2\Im(z_0)},$$

i.e., $\frac{|\mathrm{d}w|}{\Im(w)} = \frac{|\,\mathrm{d}z|}{\Im(z)}$ with $T(z) = w$. Or equivalently for every $\gamma \in SL_2(\mathbb{R})$, $(\Im\gamma z)^{-1}|\mathrm{d}\gamma z| = (\Im z)^{-1}\,\mathrm{d}z$ which shows that the differential $\mathrm{d}s^2 = y^{-2}(\,\mathrm{d}x^2 + \,\mathrm{d}y^2)$ on $\mathbb{H}$ is invariant under the group $SL_2(\mathbb{R})$. The differential generates a metric on $\mathbb{H}$. Let $\eta : [0,1] \to \mathbb{H}$ with $\eta(t) = x(t) + iy(t)$ be a smooth curve joining z with w. Then a length function can be introduced by the formula

$$\ell(\eta) = \int_0^1 \left(x'(t)^2 + y'(t)^2\right)^{\frac{1}{2}} y(t)^{-1}\,\mathrm{d}t.$$

The distance function on $\mathbb{H}$ is $d(z,w) = \inf_\eta \ell(\eta)$ where η ranges over smooth curves in $\mathbb{H}$ joining z with w. The metric derived from the differential this way is called hyperbolic metric and with it the half-plane becomes a Riemannian manifold. As it has been showed, the group $SL_2(\mathbb{R})$ acts by isometries on $\mathbb{H}$. In fact $\mathrm{Iso}^+(\mathbb{H}) \cong PSL_2(\mathbb{R})$. A path η from z to w is a geodesic if it locally minimises distances. Let $z, w \in \mathbb{H}$. If $\Re z = \Re w$, then the geodesic from z to w is the vertical line from z to w, otherwise it is the arc of circle with centre on $\mathbb{R} \cup \infty = \partial\mathbb{H}$ joining z to w. The hyperbolic metric defines the original topology on the upper half-plane. $\mathbb{H}$ equipped with the above metric is the Poincaré model of the hyperbolic plane. The hyperbolic measure on $\mathbb{H}$ can be expressed as $\mathrm{d}\mu(z) = y^{-2}\,\mathrm{d}x\,\mathrm{d}y$ in terms of the Lebesgue measure.

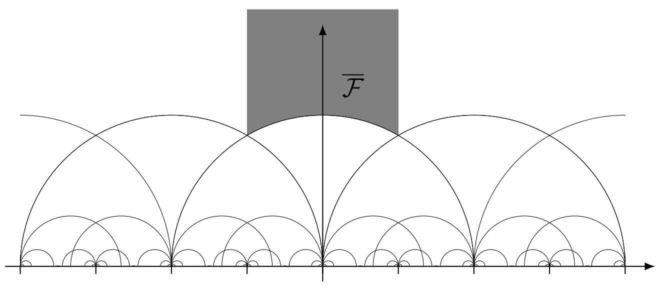
The actions of $PSL_2(\mathbb{R})$ are rigid motions of $\mathbb{H}$. For a $\gamma \in PSL_2(\mathbb{R})$, its conjugacy class will be denoted by $\{\gamma\} = \{\tau^{-1}\gamma\tau \mid \tau \in PSL_2(\mathbb{R})\}$. Conjugate motions act on $\mathbb{H}$ similarly. The identity motion forms a class by itself. Any other motion $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ has one or two fixed points in $\mathbb{C} \cup \{\infty\}$. There are three cases: if γ has one fixed point on $\mathbb{R} \cup \{\infty\}$, i.e. $|\,\mathrm{Tr}(\gamma)| = 2$, then it is called parabolic motion. A parabolic motion moves points along horocycles. If γ has two fixed points on $\mathbb{R} \cup \{\infty\}$, i.e. $|\operatorname{Tr}(\gamma)| > 2$, it is called hyperbolic motion. In this case, γ moves the points along hypercycles. Finally, an elliptic motion which moves points along circles centred at its fixed point has one fixed point in $\mathbb{H}$ and one in $\overline{\mathbb{H}}$.

As stated before, a Riemann surface of hyperbolic type is isomorphic to a quotient $\Gamma\backslash\mathbb{H}$ of $\mathbb{H}$ by some group $\Gamma \leq PSL_2(\mathbb{R})$ that acts freely and properly discontinuously. A subgroup $\Gamma \leq PSL_2(\mathbb{R})$ acting properly discontinuously on $\mathbb{H}$ is called a Fuchsian group. Poincaré showed that a subgroup of $SL_2(\mathbb{R})$ is discrete if and only if it acts properly discontinuously on $\mathbb{H}$ as a subgroup of $PSL_2(\mathbb{R})$. If Γ acts on $\mathbb{H}$, the stability group of a point $z \in \mathbb{H}$ can be defined as the subgroup $\Gamma_z = \{\gamma \in \Gamma \mid \gamma z = z\}$. If Γ is a Fuchsian group then for any $z \in \mathbb{C} \cup \{\infty\}$ the stability group Γ_z is cyclic, but not necessarily finite for $z \in \mathbb{R} \cup \{\infty\}$. A Fuchsian group is said to be of the first kind if every point on the boundary $\partial\mathbb{H}$ is a limit of an orbit Γz for some z. A Fuchsian group can be modelled by its fundamental domain in $\mathbb{H}$. A subset $\mathcal{F}_\Gamma \subset \mathbb{H}$ is called a fundamental domain for Γ if $\mathcal{F}_\Gamma$ is domain in $\mathbb{H}$, distinct points in $\mathcal{F}_\Gamma$ are not Γ-equivalent and any orbit of Γ contains a point in the closure. Every Fuchsian group of the first kind has a finite number of generators and a fundamental domain of finite area. Furthermore, a Fuchsian group of the first kind has compact fundamental domain if and only if it has no parabolic elements. In the next section specific examples of Fuchsian groups will be investigated.

3.1.3 The modular group

The group $SL_2(\mathbb{Z}) = \left\{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \mid a,b,c,d \in \mathbb{Z},\ ad - bc = 1\right\}$ which acts on $\mathbb{H}$ by the fractional linear transformations is also called modular group. The modular group is a Fuchsian group since it acts properly discontinuously on $\mathbb{H}$. $SL_2(\mathbb{Z})$ is generated by $T = \left(\begin{smallmatrix} 1 & 1 \\ & 1 \end{smallmatrix}\right)$ and $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$.

Recall that if $\mathbb{H}$ is equipped with a left action of a group Γ, then two elements of $z, w \in \mathbb{H}$ are Γ-equivalent if $z = \gamma w$ for some $\gamma \in \Gamma$. This defines an equivalence relation on $\mathbb{H}$. The Γ equivalence class of z, i.e., the orbit of z, is denoted by Γz. $\Gamma\backslash\mathbb{H}$ denotes the set of Γ-equivalence classes with a natural map $\mathbb{H} \to \Gamma\backslash\mathbb{H}$ sending $z \mapsto \Gamma z$. The set $\mathcal{F} = \{z \in \mathbb{H} \mid |\Re z| < \frac{1}{2}, |z| > 1\}$ is a fundamental domain for the modular group $\Gamma = SL_2(\mathbb{Z})$. In this case if $z, w \in \overline{\mathcal{F}}$ are $SL_2(\mathbb{Z})$-equivalent and $z \neq w$, then $z, w \in \partial\mathcal{F}$ and $w = -\overline{z}$.



Definition 3.13 (Principal congruence subgroup of $SL_2(\mathbb{Z})$)**.** Let q be a positive integer. The principal congruence subgroup of level q is

$$\Gamma(q) = \left\{\gamma \in SL_2(\mathbb{Z}) \;\middle|\; \gamma \equiv \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} \pmod{q}\right\},$$

where the matrix congruence is interpreted entrywise.

In particular $\Gamma(1) = SL_2(\mathbb{Z})$. Being the kernel of the natural homomorphism $SL_2(\mathbb{Z}) \to SL_2(\mathbb{Z}/q\mathbb{Z})$, the subgroup $\Gamma(q)$ is normal in $SL_2(\mathbb{Z})$. In fact, $\left|SL_2(\mathbb{Z}) : \Gamma(q)\right| = q^3 \prod_{p|q}\left(1 - \frac{1}{p^2}\right)$. This can seen by considering the exact sequence $1 \to \Gamma(q) \to SL_2(\mathbb{Z}) \to SL_2(\mathbb{Z}/q\mathbb{Z}) \to 1$. The surjectivity of the map $SL_2(\mathbb{Z}) \to SL_2(\mathbb{Z}/q\mathbb{Z})$ can be seen by applying the Chinese Remainder Theorem. The natural group homomorphism $SL_2(\mathbb{Z}/q\mathbb{Z}) \to SL_2(\mathbb{Z}/p_1^{\alpha_1}\mathbb{Z}) \times \cdots \times SL_2(\mathbb{Z}/p_r^{\alpha_r}\mathbb{Z})$ is bijective where $q = p_1^{\alpha_1} \ldots p_r^{\alpha_r}$. If $p > 0$ is a prime and if $\alpha \in \mathbb{Z}^+$, then $|SL_2(\mathbb{Z}/p^\alpha\mathbb{Z}| = p^{3\alpha}(1 - 1/p^2)$. As a consequence of the mentioned facts,

$$\left|SL_2(\mathbb{Z}) : \Gamma(q)\right| = |SL_2(\mathbb{Z}/q\mathbb{Z})| = \prod_{j=1}^{r} p_j^{3\alpha_j}\left(1 - \tfrac{1}{p_j^2}\right).$$

Definition 3.14 (Congruence subgroups)**.** A subgroup Γ of $SL_2(\mathbb{Z})$ is a congruence subgroup if $\Gamma(q) \subset \Gamma$ for some $q \in \mathbb{Z}^+$, in which case Γ is a congruence subgroup of level q.

Thus every congruence subgroup Γ has finite index in $SL_2(\mathbb{Z})$. Although all congruence subgroups are of finite index in $SL_2(\mathbb{Z})$, it is not true that every finite index subgroup of $SL_2(\mathbb{Z})$ is a congruence subgroup. Schreier's lemma implies that every congruence subgroup is finitely generated since $SL_2(\mathbb{Z})$ is finitely generated. Besides the principal congruence subgroups, the most important congruence subgroups are

$$\Gamma_0(q) = \left\{\gamma \in SL_2(\mathbb{Z}) \;\middle|\; \gamma \equiv \begin{pmatrix} * & * \\ & * \end{pmatrix} \pmod q\right\}$$

and

$$\Gamma_1(q) = \left\{\gamma \in SL_2(\mathbb{Z}) \;\middle|\; \gamma \equiv \begin{pmatrix} 1 & * \\ & 1 \end{pmatrix} \pmod q\right\}.$$

These groups satisfy the relation $\Gamma(q) \subset \Gamma_1(q) \subset \Gamma_0(q) \subset SL_2(\mathbb{Z})$. The map $\Gamma_1(q) \to \mathbb{Z}/q\mathbb{Z}$, $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \to b (\mathrm{mod}\ q)$ is a surjection with kernel $\Gamma(q)$. Therefore $\Gamma(q) \lhd \Gamma_1(q)$ and $\left|\Gamma_1(q) : \Gamma(q)\right| = q$. Similarly the map $\Gamma_0(q) \to (\mathbb{Z}/q\mathbb{Z})^*$, $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \to d (\mathrm{mod}\ q)$ is a surjection with kernel $\Gamma_1(q)$ and $\left|\Gamma_0(q) : \Gamma_1(q)\right| = \phi(q)$, where ϕ is the Euler totient function. It is immediate to see that $\left|SL_2(\mathbb{Z}) : \Gamma_0(q)\right| = q\prod_{p|q}\left(1 + \frac{1}{p}\right)$.

Let $\mathcal{F}$ be any fundamental domain for $SL_2(\mathbb{Z})$. Since every congruence subgroup is of finite index, a (not necessarily connected) fundamental domain for the Γ can be constructed as $\mathcal{F}_\Gamma = \bigcup_{j=1}^n \gamma_j\mathcal{F}$ where $\{\gamma_j\}_{1\le j\le n}$ is a set of coset representatives, i.e. $SL_2(\mathbb{Z}) = \bigcup_{j=1}^n \Gamma\gamma_j$.

For any congruence subgroup Γ of $SL_2(\mathbb{Z})$ acting on the upper half plane $\mathbb{H}$ one can define the corresponding modular curve $Y(\Gamma)$ as the quotient space of orbits under Γ, $Y(\Gamma) = \Gamma\backslash\mathbb{H} = \{\Gamma z \mid z \in \mathbb{H}\}$. Let $\pi_\Gamma : \mathbb{H} \to Y(\Gamma)$ denote the natural projection. $Y(\Gamma)$ is equipped with the quotient topology defined by π_Γ: a nonempty subset $U \subset Y(\Gamma)$ is open for the quotient topology if and only if $\pi_\Gamma^{-1}(U)$ is open in $\mathbb{H}$. This topology makes π_Γ a continuous function. For all nonempty open subsets $V \subset \mathbb{H}$ we have $\pi_\Gamma^{-1}\big(\pi_\Gamma(V)\big) = \bigcup_{\gamma\in\Gamma}\gamma(V)$. This set is open in $\mathbb{H}$ since every $\gamma \in \Gamma$ acts by holomorphic maps on $\mathbb{H}$ which means that π_Γ is an open mapping. $Y(\Gamma)$ is connected because $\mathbb{H}$ is also connected and π_Γ is continuous. As a consequence of the properly discontinuousness of the action of $SL_2(\mathbb{Z})$, $Y(\Gamma)$ is Hausdorff. $Y(\Gamma)$ can be made into a Riemann surface that can be compactified. At a point $\pi_\Gamma(z)$ where $z \in \mathbb{H}$ is fixed only by the identity transformation in Γ, there is a small enough neighbourhood U of z with no Γ-equivalent points in

$\mathbb{H}$ such that it is homeomorphic under π_Γ to its image $\pi(U)$ in $Y(\Gamma)$. A local inverse $\pi_\Gamma(U) \to U$ could serve as a local coordinate map.

Definition 3.15 (Elliptic points)**.** Let Γ be a congruence subgroup of $SL_2(\mathbb{Z})$. A point $z \in \mathbb{H}$ is an elliptic point for Γ if Γ_z is nontrivial as a group of transformations. The corresponding point $\pi_\Gamma(z) \in Y(\Gamma)$ is also called elliptic.

Γ_z is nontrivial as a group of transformations if the containment $\{\pm I\}\Gamma_z \supset \{\pm I\}$ of matrix groups is proper. Since $SL_2(\mathbb{Z})$ is Fuchsian group, Γ_z is finite cyclic group for every $z \in \mathbb{H}$. Thus $z \in \mathbb{H}$ has an associated positive integer $h_z = \big|\{\pm I\}\Gamma_z : \{\pm I\}\big|$. Clearly $h_z = |\Gamma_z|/2$ if $-I \in \Gamma_z$ and $h_z = |\Gamma_z|$ if $-I \notin \Gamma_z$. h_z is called the period of z under the action of Γ, with $h_z > 1$ only for the elliptic points. Let $U_z \subset \mathbb{H}$ sufficiently small so that $\gamma U_z \cap U_z = \emptyset$ for all $\gamma \in \Gamma \setminus \Gamma_z$ and $\gamma U_z = U_z$ for all $\gamma \in \Gamma_z$. Let $\tau_z \in SL_2(\mathbb{C})$ be a transformation which maps z to 0 and transforms U_z onto the unit disc $U = \{z \in \mathbb{C} \mid |z| < 1\}$. Let $e_{h_z} : U \to U$ be the power map defined by $e_{h_z} = z^{h(z)}$. Then the charts $\big(\pi_\Gamma U_z, e_m \circ \tau_z \circ \pi_\Gamma^{-1}\big)$ with z ranging over $\mathbb{H}$ and U_z form an analytic atlas for $\Gamma\backslash\mathbb{H}$.

Let $\mathbb{P}^1(\mathbb{Q}) = \mathbb{Q} \cup \{\infty\}$ be the projective line over $\mathbb{Q}$. Furthermore, let $\mathbb{H}^* = \mathbb{H} \cup \mathbb{P}^1(\mathbb{Q})$ the extended upper half-plane. The topology on $\mathbb{H}$ can be extended to $\mathbb{H}^*$ by defining a basis of open neighbourhoods of ∞ to be all sets of form $\{z \in \mathbb{H} \mid \Im z > \delta\} \cup \{\infty\}$ for $\delta > 0$, and if $r \in \mathbb{Q}$, then a basis of open neighbourhoods of r is all the sets of the form $D \cup \{r\}$ where D is any open disk in $\mathbb{H}$ whose boundary is tangent to $\Im z = 0$ axis at r. The action of $SL_2(\mathbb{Z})$ on $\mathbb{H}$ naturally extends to an action on $\mathbb{H}^*$, in which the subsets $\mathbb{H}$ and $\mathbb{P}^1(\mathbb{Q})$ are both $SL_2(\mathbb{Z})$-stable.

Definition 3.16 (Cusps)**.** If $\Gamma \leq SL_2(\mathbb{Z})$ is a congruence subgroup, then a Γ-equivalence class of points in $\mathbb{P}^1(\mathbb{Q})$ is called a cusp for Γ.

Concerning the modular group, $SL_2(\mathbb{Z})\backslash\mathbb{H}$ has one cusp, ∞, with corresponding stabilizer $\Gamma_\infty = \big\{\pm \left(\begin{smallmatrix} 1 & n \\ & 1 \end{smallmatrix}\right) \in SL_2(\mathbb{Z}) \bigm| n \in \mathbb{Z}\big\}$. Clearly ∞ is $SL_2(\mathbb{Z})$-equivalent to itself. If $r = 0$, then $0 = \left(\begin{smallmatrix} & -1 \\ 1 & \end{smallmatrix}\right)(\infty)$. If $r = a/c$ with $a, c \in \mathbb{Z} \setminus \{0\}$ and $(a, c) = 1$, one can find integers $b, d \in \mathbb{Z}$ such that $ad - bc = 1$ and $r = \gamma\infty$ where $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. If $\Gamma \subset SL_2(\mathbb{Z})$ is a congruence subgroup, since $SL_2(\mathbb{Z})$ acts transitively on the set $\mathbb{P}^1(\mathbb{Q})$ then the number of cusps for Γ is at most $|SL_2(\mathbb{Z}) : \Gamma|$. To compactify the modular curve $Y(\Gamma) = \Gamma\backslash\mathbb{H}$, let us define the extended quotient $X(\Gamma) = \Gamma\backslash\mathbb{H}^* = Y(\Gamma) \cup \big(\Gamma\backslash\mathbb{P}^1(\mathbb{Q})\big)$. The points Γr in $\Gamma\backslash\mathbb{P}^1(\mathbb{Q})$ are also called the cusps of $X(\Gamma)$. $X(\Gamma)$ can be equipped by the quotient topology inherited from the extended upper half-plane $\mathbb{H}^*$. This is a manifold. The charts can be defined similarly as in the case of $Y(\Gamma)$ except for the cusps. Let us consider a cusp $\mathfrak{a}$ for Γ. Let us consider disks $U_\mathfrak{a} \subset \mathbb{H}$ tangent to $\mathbb{P}^1(\mathbb{R})$ at $\mathfrak{a}$ sufficiently small so that $\gamma U_\mathfrak{a} \cap U_\mathfrak{a} = \emptyset$ for all $\gamma \in \Gamma \setminus \Gamma_\mathfrak{a}$ and $\gamma U_\mathfrak{a} = U_\mathfrak{a}$ for all $\gamma \in \Gamma_\mathfrak{a}$. The stability group $\Gamma_\mathfrak{a}$ is an infinite cyclic group generated by an element denoted by $\gamma_\mathfrak{a}$. There exists $\sigma_\mathfrak{a} \in SL_2(\mathbb{R})$ such that (changing the sign if necessary) $\sigma_\mathfrak{a}\infty = \mathfrak{a}$ and $\sigma_\mathfrak{a}^{-1}\gamma_\mathfrak{a}\sigma_\mathfrak{a} = \left(\begin{smallmatrix} 1 & 1 \\ & 1 \end{smallmatrix}\right)$. In this case, $\sigma_\mathfrak{a}$ is called a scaling matrix. This means that $\sigma_\mathfrak{a}^{-1}\mathfrak{a} = \infty$ and $\sigma_\mathfrak{a}$ maps the disk $U_\mathfrak{a}$ onto the half-plane $\{z \in \mathbb{H} \mid \Im z > C\}$. If C is sufficiently large, the cuspidal zones $U_\mathfrak{a}$ are pairwise disjoint. Let $e : \mathbb{H} \mapsto \{z \in \mathbb{C} \mid |z| < 1\}$ be the exponential map defined by $e(z) = e^{2\pi i z}$. Then a chart for the cusp $\mathfrak{a}$ can be defined as $\big(\pi_\Gamma U_\mathfrak{a}, e \circ \sigma_\mathfrak{a}^{-1} \circ \pi_\Gamma^{-1}\big)$. The modular curve $X(\Gamma)$ is Hausdorff, connected, and compact. Topologically, the compact Riemann surface $X(\Gamma)$ is a sphere with g handles for some $g \in \mathbb{Z}^+$. This g is the genus of $X(\Gamma)$. Using the Riemann-Hurwitz formula the genus of the modular curve $X(\Gamma)$ can be computed, and this topological invariant contains valuable arithmetic information.

Theorem 3.17 (Genus of modular curves)**.** *Let Γ be a congruence subgroup of $SL_2(\mathbb{Z})$. Let $f : X(\Gamma) \to X(\Gamma(1))$ be natural projection with degree d. Let ε_2 and ε_3 denote the number of elliptic points of period 2 and 3 in $X(\Gamma)$, and ε_∞ the number of cusps of $X(\Gamma)$. Then the genus of $X(\Gamma)$ is*

$$g = 1 + \frac{d}{12} - \frac{\varepsilon_2}{4} - \frac{\varepsilon_3}{3} - \frac{\varepsilon_\infty}{2}.$$

The proof of this theorem can be found in [DS05]. The results presented in this section can be generalized for arbitrary Fuchsian groups ([Iwa97]).

3.2 Modular forms

To define modular forms a symmetric space, a suitable subgroup of symmetries and some further properties, usually formulated in terms of differential equations and growth conditions are necessary. The symmetric space will be $\mathbb{H}$ equipped with the action of the modular group and a congruence subgroup Γ will be considered as suitable subgroup of symmetries. The differential property will be holomorphicity. The modularity property will connect these objects by requiring the function to satisfy functional equations of type $f(\gamma z) = \nu(\gamma, z)f(z)$ for all $\gamma \in \Gamma$.

3.2.1 Classical definition

There are basically two ways of looking at modular forms, the classical way defining them as complex valued functions on the upper half-plane and the adelic approach using the more advanced concept of automorphic representation. The equivalence of these two approaches are far beyond trivial and in this whole thesis the classical description will be used which explains the current section title.

Definition 3.18 (Automorphic factor of weight k)**.** Let Γ be a Fuchsian subgroup of $SL_2(\mathbb{R})$. Let $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ be an arbitrary element of Γ. An automorphic factor of weight k is a function $\nu(\gamma, z) : \Gamma \times \mathbb{H} \to \mathbb{C}$ which satisfies the following conditions:

(i) For fixed $\gamma \in \Gamma$ the function $\nu(\gamma, z)$ is holomorphic function of $z \in \mathbb{H}$;
(ii) For all $z \in \mathbb{H}$ and $\gamma \in \Gamma$ $|\nu(\gamma, z)| = |(cz + d)^k|$ for a fixed real number k;
(iii) For all $z \in \mathbb{H}$ and $\gamma, \delta \in \Gamma$, the function satisfies the chain rule $\nu(\gamma\delta z) = \nu(\gamma, \delta z)\nu(\delta, z)$;
(iv) If $-I \in \Gamma$ then for all $z \in \mathbb{H}$ and $\gamma \in \Gamma$, $\nu(-\gamma, z) = \nu(\gamma, z)$.

Every automorphic factor may be written as $\nu(\gamma, z) = \vartheta(\gamma)(cz+d)^k = \vartheta(\gamma)j_\gamma(z)$ with $|\vartheta(\gamma)| = 1$. The function $\vartheta : \Gamma \to \mathbb{C}$ is called a multiplier system. Obviously $\vartheta(I) = 1$ and $\vartheta(-I) = e^{-i\pi k}$ which is called the consistency condition. In fact, a multiplier system of weight $k \in \mathbb{Z}$ is just a unitary character of Γ which satisfies the consistency condition, namely $\vartheta(-1) = (-1)^k$. For example if $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(q)$, then $\vartheta(\gamma) = \chi(d)$ is a suitable multiplier system where χ is a Dirichlet character to modulus q. For any k, we define the weight k right action of $SL_2(\mathbb{R})$ on the set of functions $f : \mathbb{H} \to \mathbb{C}$ as follows: for any $\gamma \in SL_2(\mathbb{R})$ $f_{|\gamma}(z) = j_\gamma(z)^{-k}f(\gamma z)$, where $|\gamma$

is called the slash operator . Let f be a holomorphic function on $\mathbb{H}$ satisfying $f_{|\gamma} = \vartheta(\gamma)f$ for every $\gamma \in \Gamma$. Suppose $\mathfrak{a}$ is a cusp for Γ and $\sigma_{\mathfrak{a}}$ is a scaling matrix, i.e., $\sigma_{\mathfrak{a}}\infty = \mathfrak{a}$ and $\sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{a}}$ is the group of integral translations generated by $\left(\begin{smallmatrix}1 & 1\\ & 1\end{smallmatrix}\right)$, together with $\left(\begin{smallmatrix}-1 & -1\\ & -1\end{smallmatrix}\right)$ if $-I \in \Gamma$. Then $\gamma_{\mathfrak{a}} = \sigma_{\mathfrak{a}}\left(\begin{smallmatrix}1 & 1\\ & 1\end{smallmatrix}\right)\sigma_{\mathfrak{a}}^{-1}$ generates $\Gamma_{\mathfrak{a}}$ (with $-\gamma_{\mathfrak{a}}$ if $-I \in \Gamma$). In this setting, it can be shown that $f_{|\sigma_{\mathfrak{a}}}(\left(\begin{smallmatrix}1 & 1\\ & 1\end{smallmatrix}\right)z) = \vartheta(\gamma_{\mathfrak{a}})f_{|\sigma_{\mathfrak{a}}}(z)$. Then there exits a real number $0 \leq \kappa_{\mathfrak{a}} < 1$ such that $e(\kappa_{\mathfrak{a}}) = \vartheta(\gamma_{\mathfrak{a}})$. The functional equation for $f_{|\sigma_{\mathfrak{a}}}$ shows that the function $e(-\kappa_{\mathfrak{a}}z)f_{|\sigma_{\mathfrak{a}}}$ is periodic of period 1. Hence it can be written in the form $g(e(z))$ where $g(q)$ is a holomorphic function on $\mathbb{C}\setminus\{0\}$. f is said to be holomorhic at the cusp $\mathfrak{a}$ if $g(q)$ is holomorphic at $q = 0$. In this case the Fourier expansion of f at $\mathfrak{a}$ can be defined by expanding $g(q)$ to power series at $q = 0$.

$$f_{|\sigma_{\mathfrak{a}}}(z) = e(\kappa_{\mathfrak{a}})\sum_{n=0}^{\infty}\hat{f}_{\mathfrak{a}}(n)e(nz)$$

where the complex number $\hat{f}_{\mathfrak{a}}(n)$ are called the Fourier coefficients of f at cusp $\mathfrak{a}$. The series converges absolutely and uniformly in half-planes $\Im z \geq \varepsilon > 0$.

Definition 3.19 (Automorphic forms)**.** Let Γ be a discrete subgroup of $SL_2(\mathbb{R})$ and ϑ a multiplier system. A function $f : \mathbb{H} \to \mathbb{C}$ is said to be a (holomorphic) automorphic form for Γ of weight k with respect to the multiplier system ϑ if it satisfies the following conditions:

(i) f is holomorphic on $\mathbb{H}$;

(ii) f is holomorphic at the cups of $\Gamma\backslash\mathbb{H}$;

(iii) f satisfies the transformation rule $f_{|\gamma}(z) = j_\gamma(z)^{-k}f(\gamma z) = \vartheta(\gamma)f(z)$ for all $\gamma \in \Gamma$.

The linear space of automorphic forms for Γ with multiplier system ϑ of weight k will be denoted by $\mathcal{M}_k(\Gamma, \vartheta)$. If ϑ is trivial on $\Gamma_{\mathfrak{a}}$ for some cusp $\mathfrak{a}$, then $\mathfrak{a}$ is called a singular cusp for the multiplier system. An automorphic form $f \in \mathcal{M}_k(\Gamma, \vartheta)$ is said to be a cusp form if $\hat{f}_{\mathfrak{a}}(0) = 0$ for any singular cusp. This implies immediately that a cusp form decays exponentially at every cusp. The linear subspace of cusp forms is denoted by $\mathcal{S}_k(\Gamma, \vartheta)$.

The time has finally arrived to an example to be presented. It can be shown that the classical theta-function $\theta(z) = \sum_{n\in\mathbb{Z}} e(n^2z)$ is an automorphic form for the group $\Gamma_0(4)$ of weight $k = \frac{1}{2}$, however the multiplier system is slightly different from the trivial, precisely

$$\theta(\gamma z) = \overline{\varepsilon}_d\Big(\frac{c}{d}\Big)j_\gamma(z)^{\frac{1}{2}}\theta(z) \text{ if } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(4)$$

where

$$\overline{\varepsilon}_d = \begin{cases} 1 & \text{if } d \equiv 1 (\text{mod } 4) \\ i & \text{if } d \equiv -1 (\text{mod } 4) \end{cases}$$

and $\left(\frac{c}{d}\right)$ denotes the extended quadratic residue symbol, namely it is the Jacobi symbol if $0 < d \equiv 1 \pmod 2$ extended to all $d \equiv 1 \pmod 2$ by

$$\left(\frac{c}{d}\right) = \frac{c}{|d|}\left(\frac{c}{-d}\right) \text{ if } c \neq 0 \text{ and } \left(\frac{0}{d}\right) = \begin{cases} 1 & \text{if } d = \pm 1 \\ 0 & \text{otherwise} \end{cases}.$$

The construction of $\theta(z)$ is a special case of the following construction due to Schoenberg and Pfetzer. Let $A \in M_n(\mathbb{Z})$ be a positive definite integral matrix. Following Schimura's treatment on the material, let q be an integer such that qA^{-1} is also integral. $P(x)$ denotes a spherical harmonic relative to A, that is, a homogeneous polynomial of degree $\nu \geq 0$ for which

$$\sum_{i,j} \overline{a}_{ij} \frac{\partial^2 P}{\partial x_i \partial x_j} = 0$$

where $(\overline{a}_{ij}) = A^{-1}$. For $h \in \mathbb{Z}^n$ let

$$\theta(z, h, A, q) = \sum_{m \equiv h (\text{mod } q)} P(m) e\left(\frac{(m^T A m) z}{2q^2}\right)$$

which converges and defines a holomorphic function on $\mathbb{H}$. The functions defined this way satisfy the following transformation rules:

(i)

$$\theta(-1/z, h, A, q) = (-i)^{\nu} (\det A)^{-\frac{1}{2}} (-iz)^{\frac{n+2\nu}{2}} \sum_{\substack{k (\text{mod } q) \\ Ak \equiv 0 (\text{mod } q)}} e\left(\frac{k^T A h}{q^2}\right) \theta(z, k, A, q)$$

(ii)

$$\theta(z+2, h, A, q) = e\left(\frac{h^T A h}{q^2}\right) \theta(z, h, A, q)$$

(iii)

$$\theta(\gamma z, h, A, q) = e\left(\frac{abh^T A h}{2q^2}\right) \left(\frac{\det A}{d}\right) \left(\frac{2c}{d}\right)^n \overline{\varepsilon}_d^{-n} (cz+d)^{\frac{n+2\nu}{2}} \theta(z, ah, A, q)$$

for $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z})$ and $b \equiv 0 \pmod 2$, $c \equiv 0 \pmod{2q}$.

In particular, let $A \in M_4(\mathbb{Z})^+$. Then $Q(x) = x^T A x$ is a positive definite quaternary form. Let $r_Q(\nu) = \left|\{n \in \mathbb{Z}^4 \mid Q(n) = \nu\}\right|$. If qA^{-1} is integral, then the identities above show that

$$\theta_Q(z) = \sum_{n \in \mathbb{Z}^4} e\big((m^T A m) z\big) = \sum_{\nu=0}^{\infty} r_Q(\nu) e(\nu z)$$

is in $\mathcal{M}_2(\Gamma(q))$. This theory is elaborated in more deteil in Iwaniec's book [Iwa97].

The most important cases are when Γ is a congruence subgroup and in most cases the multiplier system will be trivial. For these forms a new terminology will be introduced.

Definition 3.20 (Modular form)**.** Let Γ be a congruence subgroup of $SL_2(\mathbb{Z})$. A modular form of weight k for Γ is a function $f : \mathbb{H} \to \mathbb{C}$ satisfying

(i) f is a holomorphic function on $\mathbb{H}$;

(ii) f is holomorphic at each cusp $\mathfrak{a}$ of $\Gamma \backslash \mathbb{H}$;

(iii) f satisfies the transformation rule $f_{|\gamma} = f$ for all $\gamma \in \Gamma$, i.e., $f(\gamma z) = (cz+d)^k f(z)$ for every $\gamma = \left(\begin{smallmatrix} * & * \\ c & d \end{smallmatrix}\right) \in \Gamma$ and every $z \in \mathbb{H}$.

The condition $f(\gamma z) = j_\gamma(z)^k f(z)$ would be impossible when k is odd and $\gamma = -I$. Hence only modular forms of even weight exists for $SL_2(\mathbb{Z})$. In the case when k is odd, either we assume that $-I \notin \Gamma$ or that we have a nontrivial multiplier system provided by an odd character χ of Γ.

3.2.2 Cusp forms and Eisenstein series

Definition 3.21 (Cusp form)**.** A modular form for a congruence subgroup Γ of $SL_2(\mathbb{Z})$ is called cusp form if for every cusp $\mathfrak{a}$ of $\Gamma \backslash \mathbb{H}^*$ the constant Fourier coefficient of f at cusp $\mathfrak{a}$ is zero.

For a congruence subgroup of $SL_2(\mathbb{Z})$ let $\mathcal{M}_k(\Gamma)$ and $\mathcal{S}_k(\Gamma)$ denote the space of modular forms and the space of cusp forms of weight k (similarly to the general case).

Let $f, g \in \mathcal{S}_k(\Gamma)$ be cusp forms of weight k. The modularity of the functions implies that $f_{|\gamma} = f$ and $g_{|\gamma} = g$ for all $\gamma \in \Gamma$. Let us take the function $\varphi(z) = f(z)\overline{g(z)}(\Im z)^k$ for $z \in \mathbb{H}$. Since

$$\begin{aligned}\varphi(\gamma z) &= f(\gamma z)\overline{g(\gamma z)}\big(\Im(\gamma z)\big)^k = f_{|\gamma}(z)j_\gamma(z)^k\overline{g_{|\gamma}(z)}\ \overline{j_\gamma(z)}^k(\Im z)^k|j_\gamma(z)|^{-2k} \\ &= f_{|\gamma}(z)\overline{g_{|\gamma}(z)}(\Im z)^k = \varphi(z)\end{aligned}$$

hence φ is Γ-invariant. Since a cusp form f has exponential decay at cusps, φ is bounded on the modular curve. This shows that in the next definition the integral is convergent.

Definition 3.22 (Petersson inner product)**.** Let Γ be a congruence subgroup of $SL_2(\mathbb{Z})$. The Petersson inner product $\langle .,.\rangle_\Gamma : \mathcal{S}_k(\Gamma) \times \mathcal{S}_k(\Gamma) \to \mathbb{C}$ is defined by

$$\langle f, g\rangle_\Gamma = \int_{\Gamma\backslash\mathbb{H}} f(z)\overline{g(z)}(\Im z)^k \,\mathrm{d}\mu(z)$$

where $\mathrm{d}\mu$ is the ($SL_2(\mathbb{Z})$-invariant) hyperbolic measure on $\mathbb{H}$.

The complex vector space $\mathcal{M}_k(\Gamma)$ of modular forms contains the subspace of cusp forms $\mathcal{S}_k(\Gamma)$. It has a complement, denoted $\mathcal{E}_k(\Gamma)$. The space $\mathcal{E}_k(\Gamma)$ consists of modular forms called Eisenstein series which do not vanish at every cusp of Γ. It can be shown that the Petersson inner product of an Eisenstein series and a cusp form is always 0, however the Petersson inner product possibly diverges for two noncusp forms. Considering the mentioned orthogonality relation, $\mathcal{M}_k(\Gamma) = \mathcal{S}_k(\Gamma) \bigoplus \mathcal{E}_k(\Gamma)$ can be written. In the further part of this section Eisenstein series will be presented. Aside from demonstrating new explicit examples of modular forms, computing the Fourier expansions of these Eisenstein series leads to natural connections between topics like Dirichlet characters, zeta and L-functions, Bernoulli numbers, theta functions and etc.

Definition 3.23 (Eisenstein series associated to ∞)**.** Let $k > 2$ and Γ a congruence subgroup of $SL_2(\mathbb{Z})$. The Eisenstein series of weight k for Γ associated to ∞ is defined by

$$E_k^{(\infty)}(z) = \sum_{\gamma \in \Gamma_\infty\backslash\Gamma} j_\gamma(z)^{-k}.$$

Definition 3.24 (Eisenstein series associated to an arbitrary cusp)**.** Let $k > 2$ and Γ a congruence subgroup of $SL_2(\mathbb{Z})$. Let $\mathfrak{a}$ be a cusp of $\Gamma\backslash\mathbb{H}^*$. Let $\sigma_\mathfrak{a}$ be a scaling matrix. The Eisenstein series of weight k for Γ associated to $\mathfrak{a}$ is

$$E_{\mathfrak{a}k}(z) = \sum_{\gamma \in \Gamma_\mathfrak{a}\backslash\Gamma} j_{\sigma_\mathfrak{a}{}^{-1}\gamma}(z)^{-k}.$$

In fact, if $E_{\mathfrak{a}k}(z)$ is an Eisenstein series associated to a cusp $\mathfrak{a}$ for a congruence subgroup Γ then it is also an Eisenstein series associated to ∞ for the congruence subgroup $\sigma_\mathfrak{a}^{-1}\Gamma\sigma_\mathfrak{a}$. Since this holds, one can assume that $\mathfrak{a} = \infty$ in most cases. If k is odd, suppose that $-I \notin \Gamma$, and $\mathfrak{a}$ is a singular cusp of Γ. $E_{\mathfrak{a}k}$ belongs to $\mathcal{E}_k(\Gamma)$ and does not vanish at $\mathfrak{a}$ while it vanishes at all other cusps of Γ. The proof of this can be found in [Sar90].

Corrollary 3.25. *Let Γ be a congruence subgroup of $SL_2(\mathbb{Z})$ with cusps $\mathfrak{a}_1, \dots, \mathfrak{a}_r$. The Eisenstein series $E_{\mathfrak{a}_j k}(z)$ span $\mathcal{E}_k(\Gamma)$ hence every $g \in \mathcal{M}_k(\Gamma)$ has a unique representation*

$$g(z) = E(z) + f(z)$$

where $E \in \mathcal{E}_k(\Gamma)$ and $f \in \mathcal{S}_k(\Gamma)$.

The Fourier series of the Eisenstein series defined above can be computed, and one finds that the n-th coefficient is an elementary arithmetical function of n and is of the form $a(n) = C\sum_{d|n} d^{k-1}F(d,n)$ where $F(d,n)$ is a periodic function in both variables (mod q) and C does not depend on n. The computation can be found in [Miy06]. This result will be derived in this section only for the modular group. Let $k > 2$ be an even integer. One can define the non-normalised Eisenstein series of weight k.

$$G_k(z) = \sum_{(c,d)\in\mathbb{Z}^2\setminus\{(0,0)\}} \frac{1}{(cz+d)^k}$$

Arranging the terms in accordance with $n = (c,d)$, $G_k(z)$ takes the form

$$G_k(z) = \left(\sum_{n=1}^{\infty} n^{-k}\right)\left(\sum_{(c,d)=1} \frac{1}{(cz+d)^k}\right) = 2\zeta(k)E_k(z).$$

Here $\zeta(k)$ is the value of the Riemann-zeta at k which is $\zeta(k) = -\frac{(2\pi i)^k}{2k!}B_k$ where B_k is the k-th Bernoulli number and

$$E_k(z) = \frac{1}{2}\sum_{\gamma\in\Gamma_\infty\setminus\Gamma} j_\gamma(z)^{-k}$$

is the Eisenstein series for $SL_2(\mathbb{Z})$ of weight k associated to ∞. Returning to G_k, the sum is absolutely convergent and converges uniformly on compact subsets of $\mathbb{H}$. It has modularity property since for a $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z})$ we have $G_k(\gamma z) = (cz+d)^k G_k(z)$. In addition, G_k is bounded as $\Im z \to \infty$ which means f is holomorphic at the only cusp of $SL_2(\mathbb{Z})$. To compute the Fourier coefficients of G_k we begin by the well-known product representation for the sine function

$$\pi z \prod_{d=1}^{\infty}\left(1 - \frac{z}{d}\right)\left(1 + \frac{z}{d}\right) = \sin(\pi z).$$

This can be established by comparing the zeros of functions standing on the two sides and applying Liouville's theorem. Taking the logarithmic derivative we get

$$\frac{1}{z} + \sum_{d=1}^{\infty}\left(\frac{1}{z-d} + \frac{1}{z+d}\right) = \pi\cot\pi z = \pi i - 2\pi i\sum_{n=0}^{\infty} e^{2\pi i n z}.$$

Differentiating $(k-1)$ times with respect to z gives

$$\sum_{d\in\mathbb{Z}} \frac{1}{(z+d)^k} = \frac{(-2\pi i)^k}{(k-1)!}\sum_{d=1}^{\infty} d^{k-1}e^{2\pi i d z}$$

for any $k \geq 2$. If $k > 2$ even, then

$$\sum_{(c,d)\in\mathbb{Z}^2\setminus\{(0,0)\}} \frac{1}{(cz+d)^k} = \sum_{n\neq 0}\frac{1}{n^k} + 2\sum_{c=1}^{\infty}\left(\sum_{d\in\mathbb{Z}} \frac{1}{(cz+d)^k}\right).$$

Or equivalently

$$\sum_{(c,d)\in\mathbb{Z}^2\setminus\{(0,0)\}} \frac{1}{(cz+d)^k} = 2\zeta(k) + 2\frac{(2\pi i)^k}{(k-1)!}\sum_{c=1}^{\infty}\sum_{d\in\mathbb{Z}} d^{k-1}e(cdz).$$

Hence by collecting the terms $cd = n$ one can derive the Fourier expansion of G_k as

$$G_k(z) = 2\zeta(k) + 2\frac{(2\pi i)^k}{\Gamma(k)}\sum_{n=1}^{\infty}\sigma_{k-1}(n)e(nz)$$

where the coefficient $\sigma_{k-1}(n)$ is the arithmetic function defined by $\sigma_s(n) = \sum_{d|n} d^s$. For the normalized Eisenstein series of weight k we get

$$E_k(z) = \frac{G_k(z)}{2\zeta(k)} = 1 - \frac{2k}{B_k}\sum_{n=1}^{\infty}\sigma_{k-1}(n)e(nz).$$

The reason for defining Eisenstein series only for $k > 2$ is that the infinite sum in the definition of G_k does not converge absolutely if $k = 2$ and hence the transformation rule $G_2(-1/z) = z^2G_2(z)$ cannot be derived. This is also true more generally because $E_2^{(\infty)}(z)$ in Definition 3.23 does not converge. Although if we define G_2 as

$$G_2(z) = \sum_{c=-\infty}^{\infty}\sum_{\substack{d=-\infty \\ (c,d)\neq(0,0)}}^{\infty}\frac{1}{(cz+d)^2} = 2\zeta(2) + \sum_{c=1}^{\infty}\sum_{d=-\infty}^{\infty}\frac{1}{(cz+d)^2} = 2\zeta(2)E_2(z),$$

then $G_2(z)$ (and $E_2(z)$) converges conditionally and holomorphic on $\mathbb{H}$. From this point a calculation can be made similar to that has been done before. So we have

$$E_2(z) = 1 - 24\sum_{n=1}^{\infty}\sigma(n)e(nz).$$

This shows that $E_2(z)$ is also holomorphic at ∞. This function is sometimes referred as quasi-modular form since it satisfies transformation law similar to required from a modular form.

Proposition 3.26. *For $\gamma = \left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right) \in SL_2(\mathbb{Z})$, the function $E_2(z)$ satisfies the transformation rules*

(i) $E_2(z+1) = E_2(z)$.

(ii) $z^{-2}E_2(-1/z) = E_2(z) + \frac{12}{2\pi i z}$.

(iii) $(cz+d)^{-2}E_2(\gamma z) = E_2(z) + \dfrac{12c}{2\pi i(cz+d)}$.

The proof of this statement can be found in [Kob12]. Although $E_2(z)$ is not a modular form it can be used to define modular forms of weight 2 for congruence subgroups of higher level as follows. For every positive integer q we define a holomorphic function $E_2^{(q)} : \mathbb{H} \to \mathbb{C}$ by

$$E_2^{(q)}(z) = E_2(z) - qE_2(qz).$$

For a $\gamma = \left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right) \in \Gamma_0(q)$ we have

$$\begin{aligned} j_\gamma(z)^{-2}E_2^{(q)}(\gamma z) &= j_\gamma(z)^{-2}E_2(\gamma z) - qj_\gamma(z)^{-2}E_2(q\gamma z) \\ &= (cz+d)^{-2}E_2\Big(\frac{az+b}{cz+d}\Big) - q\Big(\big(\tfrac{c}{q}\big)(qz)+d\Big)^{-2}E_2\Big(\frac{a(qz)+bq}{(c/q)(qz)+d}\Big). \end{aligned}$$

By applying the transformation rule of $E_2(z)$ established in Proposition 3.26, we derive

$$\begin{aligned} (cz+d)^{-2}E_2^{(q)}\Big(\frac{az+b}{cz+d}\Big) &= E_2(z) - \frac{12c}{2\pi i(cz+d)} - q\left(E_2(qz) - \frac{12\big(\frac{c}{q}\big)}{2\pi i\big((c/q)(qz)+d\big)}\right) \\ &= E_2(z) - qE_2(qz) = E_2^{(q)}(z). \end{aligned}$$

If Γ is a congruence subgroup of $SL_2(\mathbb{Z})$ then one can construct similarly $E_{\mathfrak{a}_j 2}(z)$ for each cusp $\mathfrak{a}_j$, $j = 1, \dots, r$ of Γ. The suitable combinations of these functions provide elements of $\mathcal{M}_2(\Gamma)$.

In this way one gets an $r-1$ dimensional space $\mathcal{E}_2(\Gamma)$. The Fourier coefficients of a member $E(z) \in \mathcal{E}_2(\Gamma)$ are again divisor sums and again every $g \in \mathcal{M}_2(\Gamma)$ can be uniquely expressed as $g(z) = E(z) + f(z)$ with $E(z) \in \mathcal{E}_2(\Gamma)$, $\mathcal{S}_2(\Gamma)$.

3.2.3 Hecke operators

Hecke (1937) introduced a certain ring of operators acting on modular forms. Generally, the Hecke operators are averaging operators over a suitable finite collection of double cosets with respect to a group.

Definition 3.27. Let G be a group and $\Gamma \leq G$. The commensurator $\mathrm{Comm}_G(\Gamma)$ of Γ is a subgroup of G containing Γ and defined by

$$\mathrm{Comm}_G(\Gamma) = \{g \in G \mid g^{-1}\Gamma g \cap \Gamma \text{ has finite index in both } \Gamma \text{ and } g^{-1}\Gamma g\}.$$

Let us assume that G acts on a space X. Then the Hecke operators can be defined generally on $L^2(\Gamma\backslash X)$. For $g \in \mathrm{Comm}_G(\Gamma)$ one can define the operator $T_g : L^2(\Gamma\backslash X) \to L^2(\Gamma\backslash X)$ as follows: write $\Gamma = \bigcup_{i\in I} \Delta\delta_i$ where $\Delta = g^{-1}\Gamma g \cap \Gamma$ as in the definition, then $T_g f(x) = \sum_{i\in I} f(g\delta_i x)$ for $x \in X$. When $G = SL_2(\mathbb{R})$ and Γ is a congruence subgroup, from the above construction the usual Hecke operators can be obtained, but in the next lines some details will be elaborated following Iwaniec's treatment in [Iwa97].

Throughout this section let k be a fixed nonnegative integer. For an arbitrary $A \in M_2^+(\mathbb{R})$ the slash operator is $f_{|A}(z) = (\det A)^{\frac{k}{2}} j_A(z)^{-k} f(Az)$. For a positive integer n we consider the set $G_n \subset GL_2^+(\mathbb{Z})$ which consists of the matrices with determinant equal to n. The collection

$$\Delta_n = \left\{ \begin{pmatrix} a & b \\ & d \end{pmatrix} \mid ad = n,\ 0 \leq b < d \right\}$$

is a finite set with number of elements $\sigma(n) = \sum_{d|n} d$. This set is in bijection with the set of sublattices of $\mathbb{Z}^2$ of index n by letting the rows define basis elements. It can be shown by a not too complicated calculation that Δ_n forms a complete set of right coset representatives of G_n modulo Γ.

$$G_n = \bigcup_{\delta\in\Delta_n} \Gamma\delta.$$

Furthermore there is a one-to-one correspondence between $\Delta_n \times \Gamma$ and $\Gamma \times \Delta_n$ since for any $\delta \in \Delta_n$ and $\gamma \in \Gamma$ there are unique $\gamma' \in \Gamma$, $\delta' \in \Delta_n$ such that $\delta\gamma = \gamma'\delta$. Let $\chi : \mathbb{Z} \to \mathbb{C}$ be a completely multiplicative function with $\chi(-1) = (-1)^k$. This induces a function on $GL_2(\mathbb{Z})$ by setting $\chi(\delta) = \overline{\chi}(a)$ if $\delta = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$.

Definition 3.28 (Hecke operators)**.** Let k be given and let $\chi : \mathbb{Z} \to \mathbb{C}$ be a completely multiplicative function with $\chi(-1) = (-1)^k$. Then for $n \in \mathbb{Z}^+$ the Hecke operator T_n on function $f : \mathbb{H} \to \mathbb{C}$ can be defined by the formula

$$T_n f = n^{\frac{k}{2}-1} \sum_{\delta\in\Delta_n} \overline{\chi}(\delta) f_{|\delta}$$

where $f_{|\delta} = (\det\delta)^{\frac{k}{2}} j_{|\delta}(z)^{-k} f(\delta z)$.

The expression in the definition can be written equivalently as

$$(T_n f)(z) = \frac{1}{n} \sum_{ad=n} \chi(a) a^k \sum_{0 \le b < d} f\left(\frac{az+b}{d}\right).$$

The action of Hecke operators on modular forms is of primary interest. As a first step, let us consider functions f which are automorphic with respect to the group Γ_∞, i.e., satisfying $f_{|\gamma} = \chi(\gamma) f$ for any $\pm \left(\begin{smallmatrix} 1 & u \\ & 1 \end{smallmatrix}\right)$ where $u \in \mathbb{Z}$. In fact, the functions satisfying this condition are exactly the periodic functions of period 1. For example if f has Fourier series $\sum_{0 \le m} a(m) e(mz)$ which converges absolutely in $\mathbb{H}$, then T_n acts on the coefficients of the series also.

$$\begin{aligned}(T_n f)(z) &= \frac{1}{n} \sum_{m=0}^{\infty} a(m) \sum_{ad=n} \chi(a) a^k \sum_{b (\bmod\ d)} e\left(m \frac{az+b}{d}\right) \\ &= \sum_{l=0}^{\infty} \sum_{ad=n} \chi(a) a^{k-1} a(dl) e(alz) = \sum_{m=0}^{\infty} \left(\sum_{ad=n, al=m} \chi(a) a^{k-1} a(dl) \right) e(mz)\end{aligned}$$

which yields that if $T_n f$ is given by the series $(T_n f)(z) = \sum_{m=0}^{\infty} a_n(m) e(mz)$, then $a_n(m) = \sum_{d|(m,n)} \chi(d) d^{k-1} a(mnd^{-2})$ where $a(.)$ is a Fourier coefficient of the original function. Let $\mathcal{T}$ denote the algebra over $\mathbb{C}$ generated by all T_n. This algebra is called Hecke algebra. One of the most important properties of the Hecke algebra is that it is a commutative algebra generated by the operators T_p for primes p. This is a corollary of the following theorem.

Proposition 3.29 (Properties of Hecke operators)**.** *Let k be given and let $\chi : \mathbb{Z} \to \mathbb{C}$ be a completely multiplicative function with $\chi(-1) = (-1)^k$ and let T_n be n-th Hecke operator.*

(i) *For any $m, n \ge 1$ we have*

$$T_m T_n = \sum_{d|(m,n)} \chi(d) d^{k-1} T_{mnd^{-2}}.$$

(ii) *The Hecke operators commute: $T_m T_n = T_n T_m$.*

(iii) *If $(m, n) = 1$, then $T_{mn} = T_m T_n$.*

(iv) *If p is a prime, then $T_{p^{\nu+1}} = T_p T_{p^\nu} - \chi(p) p^{k-1} T_{p^{\nu-1}}$.*

The proof of (i) can be found in [Iwa97]. (ii) and (iii) are immediate corollaries. (iv) is again a trivial consequence of the previous claims. Now we consider Hecke operators on the space of modular forms for $SL_2(\mathbb{Z}) = \Gamma(1)$ of weight k. We assume that k is even. Then for $f \in \mathcal{M}_k(\Gamma(1))$ we have $f_{|\gamma\delta} = f_{|\delta}$ if $\gamma \in \Gamma(1)$, $\delta \in \Delta_n$. Hence T_n can be written without ambiguity as

$$T_n f = n^{\frac{k}{2}-1} \sum_{\delta \in \Gamma(1) \backslash G_n} f_{|\delta}.$$

Proposition 3.30 (The action of Hecke operators on modular forms)**.** *The Hecke operator T_n maps modular form to a modular form and a cusp form to a cusp form, i.e.,*

$$T_n : \mathcal{M}_k(\Gamma(1)) \to \mathcal{M}_k(\Gamma(1)) \text{ and } T_n : \mathcal{S}_k(\Gamma(1)) \to \mathcal{S}_k(\Gamma(1))$$

Proof. Let $f \in \mathcal{M}_k(\Gamma(1))$, $\delta \in \Delta_n$ and $\gamma \in \Gamma(1)$. By the correspondence $\gamma\delta = \delta'\gamma'$ we have $f_{|\delta\gamma} = f_{|\gamma'\delta'} = f_{|\delta'}$ since f is a modular form. This implies

$$(T_n f)_{|\gamma} = n^{\frac{k}{2}-1} \sum_{\delta \in \Delta_n} f_{|\delta\gamma} = n^{\frac{k}{2}-1} \sum_{\delta' \in \Delta_n} f_{\delta'} = T_n f \text{ for any } \gamma \in \Gamma(1).$$

This means that $T_n f$ is really a modular form. From the expression for the Fourier coefficients of $(T_n f)(z)$, it can be seen that if f is a cusp form then $T_n f$ is also. □

Proposition 3.31. *The Hecke operators T_n acting on the space of cusp forms for the modular group are self-adjoint, i.e.,*

$$\langle T_n f, g\rangle_\Gamma = \langle f, T_n g\rangle_\Gamma \text{ for all } f, g \in \mathcal{S}_k(\Gamma(1)).$$

Theorem 3.32 (Hecke)**.** *In the space $\mathcal{S}_k(\Gamma(1))$ of cusp forms for the modular group there exists an orthonormal basis which consists of eigenfunctions of all the Hecke operators T_n.*

Proof. The space of cusp forms is a finite dimensional Hilbert space. Let $\{f_i\}_{i\le\nu}$ be an orthonormal basis of $\mathcal{S}_k(\Gamma(1))$. Then $T_n f_i$ can be written as the linear combination of functions f_j, i.e.,

$$T_n f_i = \sum_{j\le\nu} a_{ij}(T_n) f_j$$

with some $a_{ij}(T_n) \in \mathbb{C}$. Considering this fact, T_n can be represented by a matrix $M(T_n) = (a_{ij}(T_n))$. Since T_n is self-adjoint, $M(T_n)$ commutes with its adjoint or equivalently $M(T_n)$ is a normal matrix. If $\{M(T_n)\} = \mathcal{M} \subset M_\nu(\mathbb{C})$ is a commuting family of normal matrices then there exists a unitary matrix $U \in M_\nu(\mathbb{C})$ such that $U^{-1}AU$ is diagonal for every $M(T_n)$. This means that by a suitable linear transformation with a unitary matrix the system $\{f_i\}$ can be diagonalized which means they form an orthonormal basis of $\mathcal{S}_k(\Gamma(1))$ consisting of common eigenfunctions of all the Hecke operators. □

Let $f \in \mathcal{S}_k(\Gamma(1))$ be a Hecke eigenform which means $T_n f = \lambda(n) f$ for every $n \in \mathbb{Z}^+$. If f has the Fourier expansion $f(z) = \sum_{m=1}^{\infty} a(m)e(mz)$, then we obtain $\lambda(n)a(m) = \sum_{d|(m,n)} d^{k-1} a\big(mnd^{-2}\big)$ from the expression for the Fourier coefficients of $T_n f$. This shows that for a Hecke eigenform the identity $a(n) = \lambda(n)a(1)$ holds. In particular $a(1) \neq 0$, as otherwise f would vanish identically.

Although the Hecke theory has only beeb presented for the modular group, it can be generalized to congruence groups and to even more general settings.

3.3 Ramanujan-Petersson conjecture

In a remarkable article ([Ram16]), *On certain arithmetical functions* Transactions of the Cambridge Philosophical Society XXII (1916), Ramanujan considered the function

$$(2\pi)^{12} e^{2\pi iz} \prod_{n=1}^{\infty} \left(1 - e^{2\pi inz}\right)^{24} = (2\pi)^{12} \sum_{n=1}^{\infty} \tau(n) e^{2\pi inz} \text{ for } z \in \mathbb{H}.$$

The right hand side is understood as a definition for the arithmetic function $\tau(n)$ which is called Ramanujan tau function. He computed the first few values of $\tau(n)$ and made two fundamental conjectures about the function with his ingenious mathematical intuition. Both of his conjectures turned out to describe very general phenomena.

3.3.1 Ramanujan conjectures

In his paper Ramanujan intended to study the sum

$$\sum_{r,s}(n) := \sum_{i=0}^{n} \sigma_r(i)\sigma_s(n-i)$$

where $\sigma_s(n) = \sum_{d|n} d^s$ if $n \neq 0$ and $\sigma_s(0)$ is normalized to be $\frac{1}{2}\zeta(-s)$. He proved that if r and s are positive odd integers then

$$\sum_{r,s}(n) = \frac{\Gamma(r+1)\Gamma(s+1)\zeta(r+1)\zeta(s+1)}{\Gamma(r+s+2)\zeta(r+s+2)}\sigma_{r+s+1}(n) + \frac{\zeta(1-r)+\zeta(1-s)}{r+s}n\sigma_{r+s-1}(n) + E_{r,s}(n)$$

where $E_{r,s}(n)$ is the error term which is $O\big(n^{\frac{2}{3}(r+s+1)}\big)$. Then he went on to observe that if $r+s = 10, 14, 16, 18, 20$ or 24 then $E_{r,s}(n)$ involves only the constant $E_{r,s}(1)$, possibly $\sigma_{r+s-11}(n)$ and the function $\tau(n)$. When $r+s = 10$ the error terms take the form $E_{r,s}(1)\tau(n)$. He writes:

> *There is reason for supposing that* $\tau(n)$ *is of the form* $O\big(n^{11/2+\varepsilon}\big)$ *and not of the form* $o\big(n^{11/2}\big)$.

Conjecture 3.33 (Ramanujan conjectures)**.** *Let* $\tau(n)$ *be the function introduced by Ramanujan. Then* $\tau(n)$ *has the following properties:*

(1) $\tau(n)$ *satisfies the multiplicative relations*

$$\tau(p^{n+1}) = \tau(p)\tau(p^n) - p^{11}\tau(p^{n-1}) \text{ for } p \text{ prime}$$

$$\tau(n)\tau(m) = \tau(mn) \text{ if } (m,n) = 1$$

(2) $|\tau(n)| \leq n^{\frac{11}{2}} d(n)$ *where* $d(n)$ *denotes the number of divisors.*

If the first conjecture turns out to be true, then it is sufficient to show the second conjecture only for primes, namely $|\tau(p)| \leq 2p^{\frac{11}{2}}$. The first was essentially elementary and was proved by Mordell in 1917 and marked the beginning of Hecke's theory of Hecke operators. The conjectures can be reformulated more analytically.

Conjecture 3.34 (Analytical form of Ramanujan conjectures)**.** (1) *The Dirichlet-series generated by the normalized coefficients* $\lambda_\Delta(n) : \tau(n)n^{-11/2} =$ *has a product expansion*

$$\sum_{n=1}^{\infty} \frac{\lambda_\Delta(n)}{n^s} = \prod_{p \text{ prime}} \big(1 - \lambda_\Delta(p)p^{-s} + p^{-2s}\big)^{-1}$$

.

(2) *The denominator* $1 - \lambda_\Delta(p)p^{-s} + p^{-2s}$ *has zeros only on the line* $\Re s = 0$.

As already Ramanujan observed, the condition $|\tau(p)| \leq 2p^{\frac{11}{2}}$ is equivalent to the fact that the discriminant of the polynomial $1-\tau(p)X+p^{11}X^2$ is not positive. If one writes $p^{11}X^2-\tau(p)X+1 = (1-\alpha_1 X)(1-\alpha_2 X)$ then the condition is equivalent to that $|\alpha_1| = |\alpha_2| = p^{\frac{11}{2}}$. This is a formulation which will be useful later.

The vanishing condition revealed itself to be more difficult. This second conjecture was proved by Deligne in 1974. Deligne's proof with the full apparatus of algebraic geometry was a crowning achievement of mathematics regarded to be on a level with Wiles's proof of the Shimura-Taniyama-Weil conjecture which completed the proof of Fermat's Last Theorem.

3.3.2 The modular discriminant function

The modular forms form a graded ring, because if $f \in \mathcal{M}_k(\Gamma(1))$ and $g \in \mathcal{M}_l(\Gamma(1))$ then $fg \in \mathcal{M}_{k+l}(\Gamma(1))$. Cusp forms can be constructed using this fact. Taking the smallest common multiplier of 3 and 4 we can construct a cusp form $\Delta : \mathbb{H} \to \mathbb{C}$ of weight 12.

Definition 3.35 (Modular discriminant function)**.** Let

$$g_2(z) = \frac{(2\pi)^4}{12} G_4(z) = \frac{(2\pi)^4}{12} \left(1 + 240 \sum_{n=1}^{\infty} \sigma_3(n) e(nz)\right)$$

$$g_3(z) = \frac{(2\pi)^6}{216} G_6(z) = \frac{(2\pi)^6}{216} \left(1 - 504 \sum_{n=1}^{\infty} \sigma_5(n) e(nz)\right).$$

Then the $\Delta(z) = g_2(z)^3 - 27 g_3(z)^2$ is a cusp form of weight 12 and is called the modular discriminant function.

This function is called the modular discriminant. The name of the function comes from the theory of elliptic curves and functions. For $z \in \mathbb{H}$ let Λ_z be the lattice defined by $z\mathbb{Z} \oplus \mathbb{Z}$ in $\mathbb{C}$. Then one can define the Weierstrass $\wp$ function with respect to Λ_z. The functions $\wp$ and $\wp'$ satisfy the relation $(\wp'(w))^2 = 4(\wp(w))^3 - g_2(\Lambda_z)\wp(w) - g_3(\Lambda_z)$. In fact, up to constant multiple, $\Delta(z)$ is the discriminant of the cubic polynomial $p_z(x) = 4x^3 - g_2(\Lambda_z)x - g_3(\Lambda_z)$ which has distinct roots. This also implies that $\Delta(z) \neq 0$ for $z \in \mathbb{H}$. The Fourier coefficients of $\Delta(z)$ can be calculated and it turns out that the constant terms cancel out and hence $\Delta(z)$ is a cusp form of weight 12. For given k,

$$\dim \mathcal{M}_k(\Gamma(1)) = \begin{cases} \dim \mathcal{S}_k(\Gamma(1)) + 1 & \text{if there exists a non cusp form} \\ \dim \mathcal{S}_k(\Gamma(1)) & \text{otherwise} \end{cases}$$

because if there exists a modular form of weight k with nonvanishing constant Fourier coefficient, then a suitable multiple of that can be subtracted from any given modular form to obtain a cusp form. Since Eisenstein series with nonvanishing constant coefficient have been constructed for even $k > 2$, we see that $\dim \mathcal{M}(\Gamma(1)) = \dim \mathcal{S}_k(\Gamma(1)) + 1$. If $f \in \mathcal{M}_0(\Gamma(1))$ then f is bounded on the whole fundamental domain including ∞ and by the maximum principle, f is constant. Hence $\dim \mathcal{M}_0(\Gamma(1)) = 0$, i.e., $\mathcal{M}_0(\Gamma(1)) = \mathbb{C}$. Furthermore, the exact dimension of the space $\mathcal{M}_k(\Gamma(1))$ can be determined.

Proposition 3.36 (The space of modular forms for $SL_2(\mathbb{Z})$)**.** *Suppose that k is an even non-negative integer. The dimension of the space $\mathcal{M}_k(\Gamma(1))$ is given by*

$$\dim \mathcal{M}_k(\Gamma(1)) = \begin{cases} \lfloor k/12 \rfloor + 1 & \text{if } k \equiv 2 \pmod{12} \\ \lfloor k/12 \rfloor & \text{otherwise.} \end{cases}$$

The ring $\bigoplus_{k=0}^{\infty} \mathcal{M}_k(\Gamma(1))$ of modular forms is generated by E_4 and E_6.

Proof. First let $k = 2$ and let us consider the holomorphic function $F : \mathbb{H} \to \mathbb{C}$ defined by

$$F(w) = \int_i^w f(z)\, \mathrm{d}z.$$

Let $\gamma, \delta \in SL_2(\mathbb{Z})$. Then we have

$$F(\gamma i) = \int_i^{\gamma i} f(z)\ \mathrm{d}z = \int_i^{\gamma i} f_{|\delta}(z)\ \mathrm{d}z = \int_i^{\gamma i} f(\delta z)\mathrm{d}\delta(z) = \int_{\delta i}^{\delta\gamma i} f(z)\ \mathrm{d}z = F(\delta\gamma i) - F(\delta i).$$

This means that $F(\gamma i)+F(\delta i) = F(\delta\gamma i)$ thus $\phi : \gamma \to F(\gamma i)$ is a homomorphism between $SL_2(\mathbb{Z})$ and the abelian group $\mathbb{C}$. Hence ϕ factors through the abelianization $SL_2(\mathbb{Z})/[SL_2(\mathbb{Z}), SL_2(\mathbb{Z})]$. Since $SL_2(\mathbb{Z})$ is generated by two elements, the subgroup $[SL_2(\mathbb{Z}), SL_2(\mathbb{Z})]$ is of finite index. Therefore $F(\gamma i) \in \phi(SL_2(\mathbb{Z})/SL_2(\mathbb{Z})')$ for every $\gamma \in SL_2(\mathbb{Z})$. Since $\phi(SL_2(\mathbb{Z})/SL_2(\mathbb{Z})')$ is a finite set of complex values, we can define a holomorphic map $F^* : \mathcal{F}_{SL_2(Z)'} \to \mathbb{C}$ where $\mathcal{F}_{SL_2(Z)'}$ is the fundamental domain of $SL_2(\mathbb{Z})'\backslash\mathbb{H}^*$. Since this fundamental domain is compact, $F^*(z)$ attains its maximum in $\mathcal{F}_{SL_2(\mathbb{Z})}$ and by the maximum modulus principle F^* and so F is constant, which means $f(z) = 0$ for every $z \in \mathbb{H}$.

Now let us take the cases $k = 4, 6, 8$ or 10. Let $h = 6(12-k)$. Assume that $f \in \mathcal{S}_k(\Gamma(1))$. Then $E_{6(12-k)}(z)(f(z)/\Delta(z))^6$ is a modular form of weight 0 hence $E_{6(12-k)}(z) = c(f(z)/\Delta(z))^6$ for some $c \in \mathbb{C}$. Hence $E_{6(12-k)}(z)$ cannot have any zero on $\mathbb{H}$. Now $6(12-k)$ can be written as $12m$, where $d = 1, 2, 3$ or 4. Let us consider the function $\Delta^m(z)/E_{12m}(z)$ which is a modular form of weight 0 with zero of order m at ∞, which is a contradiction.

Now let us consider the case $k = 12$ and let $f(z) \in \mathcal{M}_{12}(\Gamma(1))$. $\Delta(z) \in \mathcal{S}_{12}(\Gamma(1))$, $\Delta(z) \neq 0$ for every $z \in \mathbb{H}$ and from the Fourier expansion we see that $\Delta(z)$ has a simple zero at ∞. Since $E_{12}(\infty) \neq 0$ there exists a constant $c_1 \in \mathbb{C}$ such that $f(z) - c_1E_{12}(z)$ vanishes at ∞, sot $\big(f(z) - c_1E_{12}(z)\big)/\Delta(z) \in \mathcal{M}_0(\Gamma(1))$. Since $\mathcal{M}_0(\Gamma(1))$ contains only constant functions, we have $f(z) = c_1E_{12}(z) + c_2\Delta(z)$, i.e., $\mathcal{M}_{12}(\Gamma(1)) = \mathbb{C}\Delta(z) \oplus \mathbb{C}E_{12}(z)$. Similarly can be proved by induction that $\mathcal{M}_k(\Gamma(1)) = \Delta(z)\mathcal{M}_{k-12}(\Gamma(1)) \oplus \mathbb{C}E_k(z)$. This implies the claim. □

Corrollary 3.37. *The space $\mathcal{S}_k(\Gamma(1))$ of cusp forms for $SL_2(\mathbb{Z})$ of weight 12 is one dimensional.*

Now let us define the Dedekind η function as follows

$$\eta(z) = e\left(\tfrac{z}{24}\right) \prod_{n=1}^{\infty} \left(1 - e(nz)\right).$$

Since the series $\sum_{n=1}^{\infty} \log\left(1 - e(nz)\right)$ converges absolutely and uniformly on the compact subsets of $\mathbb{H}$, η is holomorphic on $\mathbb{H}$ and satisfies the logarithmic differentiation rule.

Proposition 3.38. *The Dedekind η function satisfies the following transformation rule*

$$\eta(-1/z) = \left(\sqrt{-iz}\right)\eta(z) \text{ for every } z \in \mathbb{H}.$$

Proof. It will be shown that the logarithmic derivatives of the left and right side are equal. Therefore the two sides have to be equal up to a multiplicative constant. With substitution $z = i$ we get that the constant must be 1.

$$\begin{aligned}\frac{d}{dz}\log\left(\eta(z)\right) &= \frac{\pi i}{12} + 2\pi i \sum_{d=1}^{\infty} \frac{de(dz)}{1 - e(dz)} = \frac{\pi i}{12} + 2\pi i \sum_{d=1}^{\infty}\sum_{m=1}^{\infty} e(dmz) \\ &= \frac{\pi i}{12} + 2\pi i \sum_{m=1}^{\infty}\sum_{d=1}^{\infty} de(dmz) = \frac{\pi i}{12} + 2\pi i \sum_{n=1}^{\infty}\left(\sum_{d|n} d\right) e(nz) = \frac{\pi i}{12}E_2(z).\end{aligned}$$

From this we have

$$\frac{d}{dz}\log\left(\eta(-1/z)\right) = \frac{\pi i}{12}z^{-2}E_2(-1/z).$$

Calculating the logarithmic derivative of the right side of the transformation rule

$$\frac{d}{dz}\log\left(\left(\sqrt{-iz}\right)\eta(z)\right) = \frac{1}{2z} + \frac{\pi i}{12}E_2(z) = \frac{\pi i}{12}\left(E_2(z) + \frac{12}{2\pi i z}\right).$$

These are equal by Proposition 3.26, so the desired relation holds up to a constant. The proof can be finished by substituting $z = i$. □

Clearly the function $\eta(z)$ is periodic of period 1. In fact, $\eta(z)$ is an automorphic cusp form of weight $\frac{1}{2}$. Dedekind determined the multiplier system for $\eta(z)$. Precisely we have

$$\eta(\gamma z) = \vartheta(\gamma) j_\gamma(z)^{\frac{1}{2}}\eta(z) \text{ for every } \gamma \in SL_2(\mathbb{Z})$$

where $\vartheta(\gamma) = e(b/24)$ if $\gamma = \left(\begin{smallmatrix} 1 & b \\ & 1 \end{smallmatrix}\right)$, $\vartheta(-\gamma) = e(1/4)\vartheta(\gamma)$ for every $\gamma \in \Gamma$ and

$$\vartheta(\gamma) = e\left(\frac{a+d-3c}{24c} - \frac{1}{2}s(d,c)\right) \text{ for } \gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z}),\ c > 0.$$

Here $s(d,c)$ denotes the Dedekind sum

$$s(d,c) = \sum_{0\le n<c}\frac{n}{c}\psi\left(\frac{dn}{c}\right)$$

where $\psi(x) = x - \lfloor x \rfloor - \frac{1}{2}$. Another fascinating property of $\eta(z)$ is that it appears in the theory of unrestricted partitions. Namely $\eta(z)^{-1}$ has the Fourier expansion

$$e(z/24)\eta(z)^{-1} = \sum_{n=0}^{\infty} p(n)e(nz)$$

where $p(n)$ is the number of partitions $n = n_1 + \cdots + n_s$ of n into positive integers with order disregarded. G. H. Hardy and Ramanujan proved using their circle method that

$$p(n) \sim \frac{1}{4\sqrt{3}n}e^{\pi\sqrt{2n/3}}.$$

Considering the proved transformation rule of $\eta(z)$ and the periodicity of it, one can deduce that the function $\eta^{24}(z)$ is invariant under $z \to z+1$ and satisfies $\eta^{24}(-1/z) = z^{12}\eta^{24}(z)$. Also it is holomorphic on $\mathbb{H}$ and $\lim_{\Im z\to\infty}\eta^{24}(z) = 0$, so $\eta^{24}(z) \in \mathcal{S}_{12}(\Gamma(1))$, i.e., it is a cusp form for modular group of weight 12. As we have seen, this is a one dimensional space spanned by the modular discriminant $\Delta(z)$. Comparing the leading terms of the Fourier expansions we get the remarkable identity

$$\Delta(z) = (2\pi)^{12}\eta^{24}(z) = (2\pi)^{12}e(z)\prod_{n=1}^{\infty}(1 - e(nz))^{24}$$

where on the left side one can see the function Ramanujan investigated. Hence

$$\Delta(z) = (2\pi)^{12}\sum_{n=1}^{\infty}\tau(n)e(nz).$$

At this point we are ready to prove Ramanujan's first conjecture.

Theorem 3.39. *The coefficient function* $\tau(n)$ *of* $\Delta(z)$ *is multiplicative, i.e.,* $\tau(mn) = \tau(m)\tau(n)$ *if* $(n,m) = 1$, *and*

$$\tau(p)\tau(p^n) = \tau(p^{n+1}) + p^{11}\tau(p^{n-1}) \text{ if } n \ge 1.$$

Proof. As it was proven before, in the space $\mathcal{S}_k(\Gamma(1))$ of cusp forms for the modular group there exists an orthonormal basis which consists of eigenfunctions of all the Hecke operators T_n. Since $\mathcal{S}_{12}(\Gamma(1))$ is spanned by $\Delta(z)$, $\Delta(z)$ is automatically a simultaneous eigenfunction of all the Hecke operators, namely

$$T_n\Delta(z) = \tau(n)\Delta(z)$$

since $\lambda(n)$, the Hecke eigenvalue, satisfies the condition $\tau(n) = \lambda(n)\tau(1)$ and $\tau(1) = 1$. Since $T_{mn} = T_nT_m$ for $(n, m) = 1$, we have $T_{mn}\Delta(z) = \tau(mn)\Delta(z) = T_nT_m\Delta(z) = \tau(n)\tau(m)\Delta(z)$. The second property follows similarly. □

The fact that the multiplicative property of the Ramanujan function can be deduced from a very general theory suggests that it may hold in more general settings as well.

3.3.3 Ramanujan-Petersson conjecture

The natural generalization of Ramanujan's second conjecture on the size of $\tau(p)$ is the so-called Ramanujan-Petersson conjecture. A version of the conjecture says that the Fourier coefficients $a(n)$ of a holomorphic cusp form on a congruence subgroup satisfy $a(n) = O\big(n^{\frac{k-1}{2}+\varepsilon}\big)$ for every $\varepsilon > 0$. The first result in this topic is due to Kloosterman who proved that $a(n) = O\big(n^{\frac{k}{2}-\frac{1}{8}+\varepsilon}\big)$ in 1927. This result was improved by Salié in 1931 to $a(n) = O\big(n^{\frac{k}{2}-\frac{1}{6}+\varepsilon}\big)$. The famous estimate for Kloosterman sums discovered by A. Weil implies $a(n) = O\big(n^{\frac{k}{2}-\frac{1}{4}+\varepsilon}\big)$. The proof of this result already used the work of Weil on curves over finite fields. The final step was taken by Deligne in 1974 who succeeded in resolving Ramanujan-Petersson conjecture by proving the famous Weil conjectures using the complete apparatus of algebraic geometry.

From a different viewpoint, the Ramanujan-Petersson conjecture asserts that if $f \in \mathcal{S}_k(\Gamma(q))$ is a Hecke cuspidal eigenform of weight k for a congruence subgroup of level q then the eigenvalues $\lambda(p)$ satisfy the inequality $|\lambda(p)| \leq 2p^{\frac{k-1}{2}}$ for $(p, q) = 1$. This was first proved in the special case $k = 2$ by Eichler and Igusa. They solved this case by reducing the problem to the Riemann Hypothesis for curves over finite fields.

Suppose X is a smooth projective curve of genus g over the finite field $\mathbb{F}_{p^m}$, i.e., zero set of a collection of homogeneous polynomials in a projective space. Let N_m be the number of points on X defined over $\mathbb{F}_{p^m}$. Let us introduce the zeta function attached to X defined by the series

$$\zeta(X, T) := \exp\left(\sum_{m=1}^{\infty} \frac{N_m}{m} T^m\right).$$

The formula above *a priori* gives us a formal power series with rational coefficients. It was shown by E. Artin and F.K. Schmidt that there exist complex numbers $\alpha_1 \dots \alpha_{2g}$ such that

$$\zeta(X, T) = \frac{P(T)}{(1-T)(1-p^mT)} \in \mathbb{Q}(T)$$

where $P(T) = \prod_{j=1}^{2g}(1 - \alpha_jT)$. For a complex number s we obviously have $P(q^{-s}) = 0$ if and only if there is a j with $\alpha_jq^{-s} = 1$. In this case we have $\Re(s) = \frac{1}{2}$ if and only if $|a_j| = p^{\frac{1}{2}}$. This is the statement of the Riemann Hypothesis for curves over finite fields proved by A. Weil. He also observed that the definition of the zeta function makes sense for arbitrary varieties over

$\mathbb{F}_{p^m}$ and stated his famous conjectures. The relationship between the Riemann Hypothesis and the Ramanujan-Petersson conjecture was discovered by Eichler in 1954.

Similarly to Ramanujan's observation, if α_p and β_p are the roots of the polynomial $X^2-\lambda(p)X+p$ then the inequality $|\lambda(p)| \leq 2p^{\frac{1}{2}}$ is equivalent to $|\alpha_p| = |\beta_p| = p^{\frac{1}{2}}$. This means that the Ramanujan-Petersson conjecture follows from Weil's theorem if, for every prime p there exists a curve X over $\mathbb{F}_p$ such that α_p and β_p appear among the inverse zeroes (α_j's) of the zeta function attached to X. Eichler showed that the modular curve $X(\Gamma_0(q))$ arising from modular forms of weight two is an appropriate curve for this purpose with a suitable choice of q. Shimura elaborated the theory for more general congruence subgroups. The proof of Ramanujan-Petersson conjecture for higher weights is based on the generalized Weil conjectures about higher dimensional varieties.

Conjecture 3.40 (Weil conjectures 1949)**.** *Let X be an n-dimensional smooth projective variety over $\mathbb{F}_{p^n}$. Let $\zeta(X,T)$ be the zeta function attached to the variety defined the same way.*

1. Rationality. $\zeta(X,T) \in \mathbb{Q}(T)$. *More precisely*

$$\zeta(X,T) = \frac{P_1(T)P_3(T)\dots P_{2n-1}(T)}{P_0(T)P_2(T)\dots P_{2n}(T)}$$

with $P_0 \dots P_{2n} \in \mathbb{Z}[T]$ where $P_0(T) = 1-T$, $P_{2n} = 1-p^nT$ and for $1 \leq k \leq 2n-1$ the polynomials can be written as $P_k(T) = \prod_{j=1}^{b_k}(1-\alpha_{kj}T)$ in $\mathbb{C}[T]$ (the numbers α_{kj} are algebraic integers).

2. Functional equation. $\zeta(X,T)$ *satisfies the functional equation*

$$\zeta(X, p^{-n}T^{-1}) = \pm p^{\frac{nE}{2}}\zeta(X,T)$$

where E is the Euler-Poincaré characteristic.

3. Riemann hypothesis. *For $1 \leq k \leq 2n-1$ and j the numbers α_{kj} satisfy the condition $\left|\iota(\alpha_{kj})\right| = p^{\frac{k}{2}}$ for every embedding $\iota : \overline{\mathbb{Q}} \mapsto \mathbb{C}$. Specially with substitution $T = p^{-s}$ the zeros of the polynomial $P_k(p^{-s})$ are on the critical line $\Re(s) = \frac{k}{2}$.*

4. Betti numbers. *If X is obtained by reduction mod p of a variety Y over a number field $K \subset \mathbb{C}$ then $b_k = \deg P_k$ is equal to the k-th Betti number of the topological space $Y(\mathbb{C})$, i.e., the dimension of the k-th singular homology group of $Y(\mathbb{C})$.*

The rationality was proved by Bernard Dwork (1960), the functional equation by Alexander Grothendieck (1965), and the analogue of the Riemann Hypothesis and the last claim were proved by Pierre Deligne (1974).

The numbers α_{ij} are the eigenvalues of the Frobenius endomorphism acting on the ℓ-adic cohomology group. In this setting, for Hecke eigenforms of weight k the Ramanujan-Petersson conjecture can be formulated in the following way: let α_p, β_p be the roots of the polynomial $X^2 - \lambda(p)X + p^{k-1}$, then $|\alpha_p| = |\beta_p| = p^{\frac{k-1}{2}}$. Similarly to the previous case, the Ramanujan-Petersson conjecture was derived from the fact that there exists a smooth projective variety X such that the numbers α_p, β_p occur as Frobenius eigenvalues in an appropriate ℓ-adic cohomology group. For future purposes we formulate this result in a manner useful for us.

Theorem 3.41 (Ramanujan bound)**.** *Let $f(z) \in \mathcal{S}_k(\Gamma)$ be a cusp form of weight k for a congruence subgroup Γ with Fourier expansion $\sum_{n=1}^{\infty} a(n)e(nz)$. Then*

$$a(n) = O_\varepsilon\left(n^{\frac{k-1}{2}+\varepsilon}\right) \text{ for all } \varepsilon > 0.$$

There is a generalization of the Ramanujan-Petersson conjecture to the group $GL_n(\mathbb{R})$ and to automorphic forms called Maaß forms which are special non holomorphic forms. This conjecture is still unsolved except in some very special cases and it still occupies a central position in the net of open mathematical problems. As M. Ram Murty and V. Kumar Murty write in the article titled *The legacy of Srinivasa Ramanujan*

> *This conjecture, later called Ramanujan's conjecture, came to play a pivotal role in the towering edifice known as the Langlands program, a far-reaching program articulated by R.P. Langlands in the 1970s.*

Numerous historical additions can be found in the books [Sar90],[Lub10] and for a good overview of the relevance of the general Ramanujan-Petersson conjecture one can consult the survey [BB13].

4. Ramanujan graphs

4.1 Definition of Ramanujan graphs

In this chapter our main purpose is to relate the problem of constructing explicit expander families to the theory of automorphic forms via the concept of Ramanujan graph introduced by Lubotzky, Phillips and Sarnak in 1988 ([LPS86],[LPS88]).

4.1.1 Universal coverings

As Corollary 2.19 states, a family of d-regular increasing graphs is an expander family if and only if the spectral gaps of the graphs are bounded away from zero. Furthermore, the spectral gap provides an estimate for the expansion constant. On the other hand, the asymptotic threshold of Theorem 2.30 suggests how large the spectral gap can be. This motivates the definition of Ramanujan graphs. Ramanujan graphs are those regular graphs which satisfy the strongest asymptotic bound on their eigenvalues.

Definition 4.1 (Ramanujan graph)**.** A finite, connected, d-regular graph G is Ramanujan if for every eigenvalue μ of the adjacency operator either $\mu = \pm d$ or $|\mu| \leq 2\sqrt{d-1}$.

Let $G = (V, E)$ be a non-empty connected graph. Let us construct a new graph $\hat{G}_{v_0} = (\hat{V}, \hat{E})$ in the following way: The vertex set $\hat{V}$ is given by the set of all non-backtracking paths of length $k \geq 0$ in G starting at v_0, i.e., $\eta : \{1, \dots k\} \to V$ with $\eta(j+1) \in N(\eta(j))$ for $1 \leq j \leq k-1$ and $\eta(j+1) \neq \eta(j-1)$ for $2 \leq j \leq k-1$. The edge set $\hat{E} \subset \hat{V} \times \hat{V}$ is given by pairs $\{\eta_1, \eta_2\}$ where η has length $k \geq 0$, η_2 has length $k+1$, and the restriction of η_2 to the first k vertices is equal to η_1. There is furthermore a natural graph map $\psi : \hat{G}_{v_0} \to G$ such that for $\eta \in \hat{V}$ of length k, $\psi(\eta) = \eta(k)$ and for an edge $(\eta_1, \eta_2) \in \hat{E}$ with η_1 of length k and η_2 of length $k+1$ $\psi(e) = (\eta_1(k), \eta_2(k+1)) \in E$. The graph $\hat{G}_{v_0}$ is called the universal cover of the graph G. This is the usual covering concept from topology in the case of CW-complexes. If $d \geq 1$ is an integer and G is a finite non-empty connected d-regular graph, then for any $v_0 \in G$, the universal cover $\hat{G}_{v_0}$ of G based at v_0 is isomorphic to the infinite d-regular tree $\mathbb{T}_d$. The fact that every element of a collection of graphs has the same universal cover implies consequences for the spectrum of the graphs also as Greenberg's theorem states.

Theorem 4.2 (Greenberg)**.** *Let $\hat{G}$ be an infinite connected graph and $\mathcal{L}$ the family of finite graphs covered by $\hat{G}$. Let r_A be the spectral radius of the adjacency operator A acting on $L^2(\hat{G}, \mu)$. Then for every $\varepsilon > 0$ there exists $0 < c(\hat{G}, \varepsilon) < 1$, such that if $(V, E) = G \in \mathcal{L}$ with $|V| = n$, then at least $c(\hat{G}, \varepsilon)n$ of the eigenvalues satisfy $\mu \geq r_A - \varepsilon$.*

In the light of this theorem, one can look at Ramanujan graphs from another viewpoint. In fact, being Ramanujan means that all of the non-trivial eigenvalues are in the spectrum of the graph's universal cover. This naturally suggests a generalization of Ramanujan graphs. Let G be a graph not necessarily regular and let $\hat{G}$ its universal cover with spectral radius r_A. Then

G is said to be Ramanujan if $\mu(G)^* \leq r_A$. If G is d-regular, this coincides with the definition above.

4.1.2 Ihara zeta function

Motivated by the theory of the Selberg zeta function, Ihara constructed a graph-theoretic analogue of the zeta function. Let G be a finite d-regular graph with $d = q+1$. For every homotopy class C of closed paths, let $\ell(C)$ denote the minimal length of the representatives of class C. A class C is primitive if it is not a proper power of another class in the fundamental group of G. These primitive classes are called prime geodesic cycles.

Definition 4.3 (Ihara zeta function)**.** The Ihara zeta of the graph G is defined by the formula

$$\zeta_G(s) = \prod_C \left(1 - q^{-s\ell(C)}\right)^{-1}$$

where the product is over all prime geodesic cycles C and $\ell(C)$ is the length of C.

Proposition 4.4 (Ihara)**.** *Let G be a d-regular graph on n vertices with $k = q+1$. Then*

$$\zeta_G(s) = \left(1-u^2\right)^{-(r-1)} \det\left(I - Au + qu^2\right)^{-1}.$$

where $u = q^{-s}$, $r = \mathrm{rank} H^1(G,\mathbb{Z}) = (q-1)n/2$ and A is the adjacency matrix of G.

Corrollary 4.5 (Analytic characterization of Ramanujan graphs)**.** *Let G be a connected d-regular graph. Then G is a Ramanujan graph if and only if $\zeta_G(s)$ satisfies the analogue of Riemann hypothesis, i.e., all the poles of $\zeta_G(s)$ in $0 < \Re(s) < 1$ lie on the line $\Re(s) = \frac{1}{2}$.*

Proof. Let $\phi(z)$ be the characteristic polynomial of A, the adjacency matrix of G, i.e., $\phi(z) = \det(zI - A)$. Then by Proposition 4.4 we have

$$\zeta_G(s)^{-1} = \left(1-u^2\right)^{r-1} u^n \phi\Big(\frac{1-qu^2}{u}\Big)$$

where n is the number of vertices of G. Let us set $z = \frac{1+qu^2}{u}$. Then $qu^2 - zu + 1 = 0$ and $u = \frac{z \pm \sqrt{z^2-4q}}{2q}$. If $z = \pm(q+1)$ then $u = \pm 1, \pm\frac{1}{q}$ and hence $\Re(s) = 1$. The Ramanujan property is equivalent to the property that for any other z such that $\phi(z) = 0$, $|z| \leq 2\sqrt{q}$. If we assume that G is Ramanujan, then $|z| \leq 2\sqrt{q}$ so $u = \pm q^{-1/2}$ or u is non-real. The first case implies $\Re(s) = \frac{1}{2}$. In the second case z is still real, and hence $\frac{z\overline{u}}{\overline{u}} = \frac{\overline{u}+q|u|^2 u}{|u|^2}$ is also real, and thus $q|u|^2 = 1$ which means that $\Re(s) = \frac{1}{2}$. On the other hand, if s_0 is a pole with $\Re(s) = \frac{1}{2}$ then $q|u_0|^2 = 1$ and using the same identity as before we have $|z_0| = q|u_0 - \overline{u_0}| \leq 2\sqrt{q}$, so G is Ramanujan. □

For results in this direction one can consult [Ter10].

4.2 The Lubotzky-Phillips-Sarnak construction

In this section our purpose is to present finally one of the first constructions of Ramanujan graphs discovered by Lubotzky, Philips and Sarnak and establish the fascinating connection between the theory of expander graphs and the theory of automorphic forms via the notorious Ramanujan conjecture. In our discussion, we will follow the steps of Sarnak's book ([Sar90]) on this subject. Essentially the same construction and proof can be found in Lubotzky's book ([Lub10]) but in a different treatment. His approach is based on representation theory of algebraic groups.

4.2.1 Hurwitz quaternions

Let $\tilde{H}$ denote the ring of Hurwitz quaternions, i.e.,

$$\tilde{H} = \Big\{ \alpha = a_0 + a_1 i + a_2 j + a_3 k \;\Big|\; \begin{pmatrix} a_0 & a_1 \\ a_2 & a_3 \end{pmatrix} \in M_2(\mathbb{Z}) \cup M_2\left(\tfrac{1}{2}(2\mathbb{Z}+1)\right) \text{ and } i^2 = j^2 = k^2 = -1, ij = -ji = k, \text{etc.} \Big\}$$

and $H(\mathbb{Z}) \subset \tilde{H}$ be the ring of integral Hurwitz quaternions, i.e. $a_0, a_1, a_2, a_3 \in \mathbb{Z}$. The following classical theorem by Jacobi will be also necessary for us.

Theorem 4.6 (Jacobi)**.** *Let n be a positive integer. Then the number of integer solutions of $x_1^2 + x_2^2 + x_3^2 + x_4^2 = n$ is*

$$r_4(n) = 8 \sum_{d|n \text{ and } 4\nmid d} d, \text{ in particular, } r_4(p) = 8(p+1).$$

There are numerous proofs of this theorem, an analytic proof using θ functions can be found in [SS03]. Hurwitz also proved this theorem by using the properties of quaternions named after him. We sketch his proof briefly in the case $n = p$. $\tilde{H}$ is a principal ideal domain with a unit group consisting of 24 elements. Two elements $\alpha, \beta \in \tilde{H}$ generate the same left ideal if and only if $\alpha = \varepsilon\beta$ where ε is a unit. Hence the number of elements of norm p is equal to 24 times the number of proper left ideals containing $p\tilde{H}$ properly. Since p is odd $\tilde{H}/p\tilde{H} \cong \tilde{H}(\mathbb{Z}/p\mathbb{Z}) \cong M_2(\mathbb{F}_p)$. This implies that the number of elements of norm p is equal to 24 times the proper left ideals of $M_2(\mathbb{F}_p)$. These are principal left ideals generated by non-zero non-invertible elements. It can be shown that their number is $(p+1)(p^2-1)$. The group $GL_2(\mathbb{F}_p)$ acts on them by left multiplication, and two elements are in the same orbit if and only if they generate the same left ideal. The size of these orbits turn out to be $p^2 - 1$ and hence there are $p+1$ left ideals and $24(p+1)$ elements of norm p in $\tilde{H}$. One can check that $8(p+1)$ of them have integral entries.

Let p be an odd prime, $p \equiv 1 (\text{mod } 4)$. Let $N(\alpha)$ be the integer $\alpha\overline{\alpha} = \sum a_l^2$ for every $\alpha \in H(\mathbb{Z})$. The units of the ring $H(\mathbb{Z})$ are $\pm 1, \pm i, \pm j, \pm k$. Let S_p'' be the set of all $\alpha \in H(\mathbb{Z})$ with $N(\alpha) = p$. By Jacobi's theorem $|S_p''| = 8(p+1)$. It is clear that for such an α there is precisely one α_l which is odd. The set S_p'' is fixed under the action of the units. For each $\alpha' \in S_p''$ there is a unique associate $\alpha = \varepsilon\alpha'$ for which $N(\alpha) = p$, $\alpha \equiv 1 \pmod 2$ in $H(\mathbb{Z})$ and $a_0 > 0$. Let S_p' be the set of these $(p+1)$ representatives. This set consists of $s := (p+1)/2$ conjugate pairs, i.e.,

$$S_p' = \{\alpha_1, \overline{\alpha}_1, \ldots, \alpha_s, \overline{\alpha}_s\}.$$

A word on the elements S_p' is called reduced if there is no conjugate pair $(\alpha_l\overline{\alpha}_l)$ in it.

Lemma 4.7. *Every $\alpha \in H(\mathbb{Z})$ with $N(\alpha) = p^k$ can be expressed uniquely in the form $\alpha = \varepsilon p^r R_m(\alpha_1, \overline{\alpha}_1, \ldots, \alpha_s, \overline{\alpha}_s)$ where ε is a unit, $2r + m = k$ and R_m is a reduced word in the α_l's of length m.*

Proof. According to the result by Dickson ([Sar90]), $H(\mathbb{Z})$ is a left and right Euclidean ring and an $\alpha \in H(\mathbb{Z})$ with $N(\alpha) \equiv 1 \pmod 2$ is prime if and only if its norm is prime. We prove this by induction on k. If $k = 1$ then the claim is simply a reformulation of the previous discussion. Hence let us assume that $k > 1$. Since $N(\alpha) = p^k$, α can be written as $\alpha = \gamma\beta$ where $N(\gamma) = p^{k-1}$ and $N(\beta) = p$. By the definition of S_p, there is a unique ε such that $\alpha = \gamma\varepsilon\alpha_l$ with $\alpha_l \in S'_p$. Then by applying the inductive hypothesis $\gamma\varepsilon$ has the required form. After performing potential cancellations we arrive at $\alpha = \varepsilon p^r R_m(\alpha_1, \overline{\alpha}_1, \ldots, \alpha_s, \overline{\alpha}_s)$ for some r and m.

In order to prove uniqueness one can count the number of such representations. The number of reduced words is $(p+1)p^{m-1}$ for $m \geq 1$ and is 1 if $m = 0$. Hence the number of expressions $\varepsilon p^r R_m(\alpha_1, \overline{\alpha}_1, \ldots, \alpha_s, \overline{\alpha}_s)$ with $2r + m = k$ is

$$8\left(\sum_{0\leq r<k/2} (p+1)p^{k-2r-1} + \delta(k)\right) \text{ where } \delta(k) = \begin{cases} 1 & \text{if } k \text{ is even} \\ 0 & \text{otherwise} \end{cases}.$$

The number of expressions is:

$$8\left(\sum_{0\leq j<k/2} p^{k-2r} \sum_{0\leq j<k/2} p^{k-2r-1} + \delta(k)\right) = 8\left(\sum_{j=0}^{k} p^j\right) = 8\left(\frac{p^{k+1}-1}{p-1}\right) = 8\sum_{d|p^k} d$$

which is the number of $\alpha \in H(\mathbb{Z})$ with $N(\alpha) = p^k$, hence each of such expressions represents a distinct element. □

If $\alpha \equiv 1 \pmod 2$ and $N(\alpha) = p^k$ then since $\alpha_l \equiv 1 \pmod 2$ we see that α is uniquely expressed in the form $\alpha = \pm p^r R_m(\alpha_1, \overline{\alpha}_1, \ldots, \alpha_s, \overline{\alpha}_s)$ with $2r + m = k$. We state this as a corollary.

Corrollary 4.8. *If $\alpha \equiv 1 \pmod 2$ and $N(\alpha) = p^k$, then $\alpha = \pm p^r R_m(\alpha_1, \overline{\alpha}_1, \ldots, \alpha_s, \overline{\alpha}_s)$ with $2r + m = k$, $s = (p+1)/2$ and this representation is unique.*

For $p \equiv 1 \pmod 4$ let us consider the following subset of integral Hurwitz quaternions

$$\Lambda'(2) = \left\{\alpha \in H(\mathbb{Z}) \;\middle|\; \alpha \equiv 1 \pmod 2 \text{ and } N(\alpha) = p^\nu, \nu \in \mathbb{Z}\right\}.$$

$\Lambda'(2)$ is closed under multiplication, and if we identify α and β in $\Lambda'(2)$ whenever $\pm p^l\alpha = \beta$ for some $l \in \mathbb{Z}$, then the equivalence classes so obtained form a group with operation $[\alpha\beta] = [\alpha][\beta]$ and $[\alpha][\overline{\alpha}] = [1]$. This quotient map is denoted by π_Λ. Let us denote this group by $\Lambda(2)$. For $\alpha \in H(\mathbb{Z})$ the previous claim ensures that there exists a unique representation $\pm p^r R_m(\alpha_1, \overline{\alpha}_1, \ldots, \alpha_s, \overline{\alpha}_s)$ which can be identified as a word in $[\alpha_1], \ldots [\alpha_s]$, which exactly means that the group $\Lambda(2)$ is free.

Corrollary 4.9. *$\Lambda(2)$ is a free group on the $s = \frac{p+1}{2}$ generators*

$$[\alpha_1], [\alpha_2] \ldots [\alpha_s].$$

4.2.2 The construction

Definition 4.10 (Cayley graph)**.** The Cayley graph of a group Γ with respect to a symmetric subset $S \subset \Gamma$ (i.e., $s \in S$ implies $s^{-1} \in S$) is the graph whose vertex set is Γ, and the edges are given by $\{(\gamma, \gamma s) \mid \gamma \in \Gamma, s \in S\}$. This graph is denoted by $\mathcal{G}(\Gamma, S)$.

Let p and q be unequal primes with $p, q \equiv 1 \pmod 4$. Let u be an integer satisfying $u^2 \equiv -1 \pmod q$. From Jacobi's theorem we know that there are $8(p+1)$ solutions $\alpha = (a_0, a_1, a_2, a_3)$ to $a_0^2 + a_1^2 + a_2^2 + a_3^2 = p$. Among these there are exactly $p+1$ with $a_0 > 0$ and odd. To each such α a matrix can be associated as follows

$$\alpha \mapsto \begin{pmatrix} a_0 + ua_1 & a_2 + ua_3 \\ -a_2 + ua_3 & a_0 - ua_1 \end{pmatrix}.$$

Let $S_{p,q}$ be the set of these matrices in $PGL_2(\mathbb{Z}/q\mathbb{Z})$. Then the Cayley graphs

$$X^{p,q} = \mathcal{G}\big(PGL_2(\mathbb{Z}/q\mathbb{Z}), S_{p,q}\big)$$

are $(p+1)$-regular of order $n = q(q^2 - 1)$. If $\left(\frac{p}{q}\right) = 1$ then this graph is not connected since the elements of $S_{p,q}$ all lie in the index 2 subgroup $PSL_2(\mathbb{Z}/q\mathbb{Z})$. Hence the final definition is

$$X^{p,q} = \begin{cases} \mathcal{G}\big(PGL_2(\mathbb{Z}/q\mathbb{Z}), S_{p,q}\big) & \text{if } \left(\frac{p}{q}\right) = -1 \\ \mathcal{G}\big(PSL_2(\mathbb{Z}/q\mathbb{Z}), S_{p,q}\big) & \text{if } \left(\frac{p}{q}\right) = 1. \end{cases}$$

Theorem 4.11 (Lubotzky-Phillips-Sarnak)**.** *For p and q primes with $p \neq q$, $p, q \equiv 1 \pmod 4$ let $X^{p,q}$ be the Cayley graph defined above. Then*

(i) *if* $\left(\frac{p}{q}\right) = -1$,
  - (a) $X^{p,q}$ *is a bipartite Ramanujan graph;*
  - (b) $\text{girth}(X^{p,q}) \geq 4\log_p q - \log_p 4$;
  - (c) $\text{diam}(X^{p,q}) \leq 2\log n + 2\log_p 2 + 1$ *where* $n = q(q^2 - 1)$;

(ii) *if* $\left(\frac{p}{q}\right) = 1$,
  - (a) $X^{p,q}$ *is a Ramanujan graph;*
  - (b) $\text{girth}(X^{p,q}) \geq 2\log_p q$;
  - (c) $\text{diam}(X^{p,q}) \leq 2\log n + 2\log_p 2 + 1$ *where* $n = q(q^2 - 1)/2$;
  - (d) $\alpha(X^{p,q}) \leq \frac{2\sqrt{p}}{p+1} n$ *where* $n = q(q^2 - 1)/2$;
  - (e) $\chi(X^{p,q}) \geq \frac{p+1}{2\sqrt{p}}$.

Let $(q_m)_{m\in\mathbb{Z}^+}$ be a sequence of primes with $p \neq q_m$ and $q \equiv 1 \pmod 4$ such that $\lim_{m\to\infty} q_m = \infty$. Since $X^{p,q}$ is a $d = (p+1)$-regular graph on $n = q(q^2 - q)$ or $q(q^2 - 1)/2$ vertices depending on the sign of $\left(\frac{p}{q}\right)$, the Ramanujan graph property (a) of the collection $(X^{p,q_m})_{m\in\mathbb{Z}^+}$ implies that this is an expander family also.

It is also worth considering the construction from another viewpoint. Corollary 4.9 states that the group $\Lambda(2)$ is free with generators $S_p = \{[\alpha_1], \ldots [\alpha_s]\}$. Hence the Cayley graph $\mathcal{G}\big(\Lambda(2), S_p\big)$ is an infinite tree of degree $p+1$. In order to obtain a finite graph one should consider the quotient of $\Lambda(2)$ by a suitable normal subgroup Γ of finite index. Let Γ be such a subgroup. Then Γ acts on $\Lambda(2)$ by multiplication on the right and the Cayley graph $\mathcal{G}\big(\Lambda(2)/\Gamma, S_\Gamma \Gamma\big)$ is finite where $S_\Gamma = \{\alpha_1\Gamma, \overline{\alpha}_2\Gamma \ldots \alpha_s\Gamma, \overline{\alpha}_s\Gamma\}$. This graph is $(p+1)$-regular and connected.

Let $H(\mathbb{Z}/m\mathbb{Z})$ denote the ring of quaternions with entries in $\mathbb{Z}/m\mathbb{Z}$. Let $H(\mathbb{Z}/m\mathbb{Z})^*$ be the group of invertible elements of $H(\mathbb{Z}/m\mathbb{Z})$ under multiplication. Let $Z \leq H(\mathbb{Z}/(2q)\mathbb{Z})^*$ be the central subgroup, i.e., $Z = \{a_0 \mid a_0 \neq 0\}$. Let us consider the homomorphism $\phi_1 : \Lambda(2) \to H(\mathbb{Z}/(2q)^*/Z$ which defined by $\phi_1([\alpha]) = \big(\alpha \pmod{2\text{q}}\big)Z$. The kernel of this homomorphism is the subgroup

$$\Lambda(2q) = \big\{[\alpha] \in \Lambda(2) \;\big|\; (2q)|a_j \;\; (j = 1, 2, 3) \;\; \text{where } \alpha = a_0 + a_1 i + a_2 j + a_3 k\big\} \leq \Lambda(2).$$

Our purpose is to show that the graph $X^{p,q}$ may be identified with the Cayley graph of the group $\Lambda(2)/\Lambda(2q)$ with respect to the generators $S_{(2q)}$.

Let us consider the following diagram

$$\begin{array}{ccccccc}
S'_p \subset \Lambda(2)' & & \xrightarrow{\quad\quad\tau\quad\quad} & & H(\mathbb{F}_q)^* & \xrightarrow{\sigma} & GL_2(\mathbb{F}_q) \\
\downarrow{\scriptstyle \pi_\Lambda} & & & & \downarrow{\scriptstyle \pi_q} & & \downarrow{\scriptstyle \pi_M} \\
S_p \subset \Lambda(2) & \xrightarrow{\phi_1} & H(\mathbb{Z}/(2q)\mathbb{Z})^*/Z & \xrightarrow{\phi_2} & H(\mathbb{F}_q)^*/Z & \xrightarrow{\phi_3} & PGL_2(\mathbb{F}_q)
\end{array}$$

$$S_p \subset \Lambda(2) \xrightarrow{\quad\psi\quad} PGL_2(\mathbb{F}_q)$$

where φ as defined above, π_Λ, π_q, π_M are the natural quotient maps and the other maps will be described now. Recall that u is a fixed integer satisfying $u^2 \equiv -1 \pmod q$. Let v be another integer with $u^2 + v^2 = 1$. Let $\sigma : H(\mathbb{Z}/q\mathbb{Z})^* \to GL_2(\mathbb{Z}/q\mathbb{Z})$ be defined by

$$\sigma(a_0 + a_1 i + a_2 j + a_3 k) = \begin{pmatrix} a_0 + a_1 u + a_3 v & -a_1 v + a_2 + a_3 u \\ -a_1 v - a_2 + a_3 u & a_0 - a_1 u - a_3 v \end{pmatrix}.$$

It is not difficult to check that σ is an isomorphism and sends Z to the scalar matrices of $GL_2(\mathbb{Z}/q\mathbb{Z})$ hence $H(\mathbb{Z}/q\mathbb{Z})^* \to GL_2(\mathbb{Z}/q\mathbb{Z}) \to PGL_2(\mathbb{Z}/q\mathbb{Z})$ is an exact sequence. Let ψ : $\Lambda(2) \to PGL_2(\mathbb{Z}/q\mathbb{Z})$ be the homomorphism defined by

$$[\alpha] \mapsto \begin{pmatrix} a_0 + ua_1 & a_2 + ua_3 \\ -a_2 + ua_3 & a_0 - ua_1 \end{pmatrix}$$

where $a_0 + a_1 i + a_2 j + a_3 k \equiv \alpha \pmod q$. Now ψ factors as

$$\Lambda(2) \xrightarrow{\phi_1} H(\mathbb{Z}/(2q)\mathbb{Z})^*/Z \xrightarrow{\phi_2} H(\mathbb{Z}/q\mathbb{Z})^*/Z \xrightarrow{\phi_3} PGL_2(\mathbb{Z}/q\mathbb{Z}).$$

where ϕ_3 is an isomorphism since σ maps Z to the kernel of $\pi_M : GL_2(\mathbb{F}_q) \to PGL_2(\mathbb{F}_q)$. Finally $\tau : \Lambda(2)' \to H(\mathbb{F}_q)$ is just the reduction modulo q.

In order to prove that $X^{p,q}$ may be identified with the Cayley graph of the group $\Lambda(2)/\Lambda(2q)$ with respect to the generators S_{2q} it suffices to prove that $\Lambda(2)/\Lambda(2q) \cong PGL_2(\mathbb{Z}/q\mathbb{Z})$ or $PSL_2(\mathbb{Z}/q\mathbb{Z})$ depending on the sign of $\left(\frac{p}{q}\right)$ since the generators $S_{(2q)}$ maps to $S_{p,q}$. As we have seen, ϕ_3 is an isomorphism, and $\operatorname{Ker}\psi$ is $\Lambda(2q)$,hence for proving the claim made here is sufficient to show

$$\psi(\Lambda(2)) = \begin{cases} PGL_2(\mathbb{Z}/q\mathbb{Z}) & \text{if } \left(\frac{p}{q}\right) = -1 \\ PSL_2(\mathbb{Z}/q\mathbb{Z}) & \text{if } \left(\frac{p}{q}\right) = 1 \end{cases}$$

since $\operatorname{Im}(\psi) \cong \Lambda(2)/\operatorname{Ker}(\psi) \cong \Lambda(2)/\Lambda(2q)$.

Lemma 4.12. *If ψ is defined as above then*

$$\operatorname{Im}(\psi) = \begin{cases} PGL_2(\mathbb{Z}/q\mathbb{Z}) & \text{if } \left(\frac{p}{q}\right) = -1 \\ PSL_2(\mathbb{Z}/q\mathbb{Z}) & \text{if } \left(\frac{p}{q}\right) = 1. \end{cases}$$

Proof. For $\alpha \in H(\mathbb{Z}/q\mathbb{Z})$, $N(\alpha) = \sigma(\alpha)$, where $\sigma : H(\mathbb{Z}/q\mathbb{Z})^* \to GL_2(\mathbb{F}_p)$ is as above. It is well known that for $A \in GL_2(\mathbb{F}_q)$, the associated Möbius transformation is in $PSL_2(\mathbb{F}_p)$ if and only if $\det A$ is a square in $\mathbb{F}_p^*$. This implies that if $\alpha \in H(\mathbb{Z})$ and $N(\alpha) = p$ then $\psi([\alpha])$ lies in

$PSL_2(\mathbb{Z}/q\mathbb{Z})$ if and only if $\left(\frac{p}{q}\right) = 1$. Hence $S_{p,q} \subset PGL_2(\mathbb{Z}/q\mathbb{Z}) \setminus PSL_2(\mathbb{Z}/q\mathbb{Z})$ if $\left(\frac{p}{q}\right) = -1$, and $S_{p,q} \subset PSL_2(\mathbb{Z}/q\mathbb{Z})$ otherwise. Since $\left|PGL_2(\mathbb{Z}/q\mathbb{Z})/PSL_2(\mathbb{Z}/q\mathbb{Z})\right| = 2$ it is sufficient to show that $PSL_2(\mathbb{Z}/q\mathbb{Z}) \subset \psi(\Lambda(2))$.

In this case, ψ factors as mentioned above. In that composition ϕ_3 is an isomorphism so one can concentrate on $\pi_2 \circ \pi_1$. Then the proposition is equivalent to the condition that if $\beta = b_0 + b_1 i + b_2 j + b_3 k$ is in $H(\mathbb{Z}/q\mathbb{Z})$ and $N(\beta) \equiv 1 \pmod q$ then there is an $\alpha \in H(\mathbb{Z})$ with $N(\alpha) = p^k$, $\alpha \equiv 1 \pmod 2$ and $\alpha \equiv \beta \pmod q$. Let such a β be given. Let $\gamma = c_0 + c_1 i + c_2 j + c_3 k$ where $c_0 \equiv b_0 \pmod q$, $2c_j \equiv b_j \pmod q$, for $j = 1, 2$ and 3. Then

$$c_0^2 + 4c_1^2 + 4c_2^2 + 4c_3^2 \equiv 1 \pmod q.$$

Here some results from the theory of quadratic Diophantine equations are needed ([Sar90],[LPS88]).

Theorem 4.13 (Malyshev)**.** *Let $f(x_1, \dots, x_n)$ be a quadratic form in $n \geq 4$ variables with integral coefficients and discriminant α. Let $(g, 2d) = 1$ be such that for m sufficiently large with $(g, 2md) = 1$ and m generic for f, i.e., $f = m$ may be solved mod ℓ for every ℓ, and for $(b_1, \dots, b_n, g) = 1$, $f(b_1, \dots, b_n) \equiv m \pmod g$, there exist integers*

$$(a_1, \dots, a_n) \equiv (b_1, \dots, b_n) \pmod g \text{ such that } f(a_1, \dots a_n) = m.$$

This theorem can be applied to $f(x_1, x_2, x_3, x_4) = x_1^2 + 4x_2^4 + 4x_3^2 + 4x_4^2$, $m = p^k$, $g = q$, and $(b_0, b_1, b_2, b_3) = (c_0, c_1, c_2, c_3)$. If k is large enough and $p^k \equiv 1 \pmod q$ then $f(c_0, c_1, c_2, c_3) \equiv p^k \pmod q$ and p^k is generic for f. Hence there is an $(a_0, a_1, a_2, a_3) \equiv (c_0, c_1, c_2, c_3) \pmod q$ satisfying

$$a_0^2 + 4a_1^2 + 4a_2^2 + 4a_3^2 = p^k.$$

This implies that if $\alpha = a_0 + 2a_1 i + 2a_2 j + 2a_3 k$ then $N(\alpha) = p^k$, $\alpha \equiv 1 \pmod 2$, and $\alpha \equiv \beta \pmod q$. □

This means that $S_{p,q}$ generates either $PSL_2(\mathbb{Z}/q\mathbb{Z})$ or $PGL_2(\mathbb{Z}/q\mathbb{Z})$, according to whether $\left(\frac{p}{q}\right) = 1$ or -1. As discussed previously, this implies that

Corrollary 4.14. *The Cayley graphs obtained from group $\Lambda(2)/\Lambda(2q)$ and described at the beginning of the section are isomorphic*

$$\mathcal{G}\big(\Lambda(2)/\Lambda(2q)\big) \cong \begin{cases} \mathcal{G}\big(PGL_2(\mathbb{Z}/q\mathbb{Z}), S_{p,q}\big) & \text{if } \left(\frac{p}{q}\right) = -1 \\ \mathcal{G}\big(PSL_2(\mathbb{Z}/q\mathbb{Z}), S_{p,q}\big) & \text{if } \left(\frac{p}{q}\right) = 1. \end{cases}$$

4.2.3 Proof of the theorem

Proof. (Theorem 4.11) The connectedness of $X^{p,q}$ follows from Corollary 4.14 since a Cayley graph $\mathcal{G}(\Gamma, S)$ is connected if and only if S generates Γ and hence $\mathcal{G}\big(\Lambda(2)/\Lambda(2q)\big)$ is connected by definition (see Corollary 4.9). If there exists a homomorphism ϕ from Γ to the multiplicative group $\{1, -1\}$ such that $\chi(S) = \{-1\}$ then $\mathcal{G}(\Gamma, S)$ is bipartite. The converse holds provided $\mathcal{G}(\Gamma, S)$ is connected. This implies that if $\left(\frac{p}{q}\right) = -1$ then $X^{p,q}$ is bipartite since $PGL_2(\mathbb{Z}/q\mathbb{Z})/PSL_2(\mathbb{Z}/q\mathbb{Z}) \cong \{\pm 1\}$.

The claims (i)(c), (ii)(c), (d) and (e) follow from the Ramanujan property (see Definition 4.1) as Propositions 2.22, 2.23 and 2.24 show since in this case $\mu(X^{p,q}) \leq 2\sqrt{p}$.

Claim 4.15 (Girth lower bound). *If* $\left(\frac{p}{q}\right) = 1$ *then* $\mathrm{girth}(X^{p,q}) \geq 2\log_p q$ *and if* $\left(\frac{p}{q}\right) = -1$ *then* $\mathrm{girth}(X^{p,q}) \geq 4\log_p q - \log_p 4$.

Proof. Let $g = \mathrm{girth}(X^{p,q})$. Since $X^{p,q}$ is a Cayley graph it is vertex-transitive and hence without loss of generality it can be assumed that the shortest circuit runs from 1 to itself. On the tree $\mathcal{G}(\Lambda(2), S_p)$ g amounts to the length of the smallest member of $\Lambda(2q)$. If $\gamma \in \Lambda(2q)$, $\gamma \neq 1$, is of length g then we can find an integral quaternion $\gamma' \in \Lambda(2)'$ such that

$$\gamma' = \beta_1\beta_2 \ldots \beta_g \text{ with } \beta_j \in \{\alpha_1, \overline{\alpha}_1, \ldots \alpha_s, \overline{\alpha}_s\} \text{ and } \gamma' \in \Lambda(2)'.$$

Hence $N(\gamma') = p^g$ and $\gamma' = a_0 + 2qa_1i + 2qa_2j + 2qa_3k$, $a_l \in \mathbb{Z}$. Since γ is not equivalent to 1 in Λ', at least one of a_1, a_2, a_3 is nonzero. Thus we have

$$p^g = a_0^2 + 4q^2a_1^2 + 4q^2a_2^2 + 4q^2a_3^2.$$

If $\left(\frac{p}{q}\right) = 1$ this identity implies that $p^g \geq 4q^2$ or equivalently $g \geq 2\log_p q$. Suppose now $\left(\frac{p}{q}\right) = -1$. Since $p^g \equiv a_0^2 \pmod q$, we have $1 = \left(\frac{p^g}{q}\right) = \left(\frac{p}{q}\right)^g = (-1)^g$, so that $g = 2r$ for some $r \in \mathbb{Z}$. Returning to the congruence $p^{2r} \equiv a_0^2 \pmod{q^2}$, we have $p^r \equiv \pm a_0 \pmod{q^2}$. Assume by contradiction that $g < 4\log_p q - \log_p 4 = \log_p \frac{q^4}{4}$, so $p^r < \frac{q^2}{2}$. Then $|p^r \pm a_0| < q^2$ since $a_0^2 \leq p^g$ and so $|a_0| \leq p^r$. From this one can deduce that $p^r = \pm a_0$, but then $p^g = a_0^2$ which forces a_1, a_2, a_3 to be zero, in contradiction with our assumptions. □

To complete the proof of Theorem 4.11 the verification of Ramanujan property is still needed.

Claim 4.16 (Spectral estimates). *Let* $(p+1) = \mu_0 > \mu_1 \geq \cdots \geq \mu_{n-1}$ *denote the spectrum of the adjacency operator* $A_{X^{p,q}}$. *If the eigenvalues are written in the form* $\mu_j = 2\sqrt{p}\cos\theta_j$ *then the numbers* θ_j *are real, i.e.,* $|\mu_j| \leq 2\sqrt{p}$.

Proof. Let $\mathbb{T}_{p+1}$ be the infinite $(p+1)$-regular tree with vertex set V_{p+1}. Let Γ be a discontinuous group of automorphism acting on $\mathbb{T}_{p+1}$. Let us consider the adjacency operator A acting on $L^2(\mathbb{T}_{p+1}, \mu)$ defined by

$$Af(x) = \sum_{d(x,y)=1} f(y),$$

where d is the distance on the tree. This is an integral operator

$$Af(x) = \int_{V_{p+1}} k_A(x,y)f(x)\,\mathrm{d}\mu(x) \text{ with } k_A(x,y) = \begin{cases} 1 & \text{if } d(x,y) = 1 \\ 0 & \text{if } d(x,y) \neq 1. \end{cases}$$

where $k_A(x,y)$ is the kernel of the operator . Let A_n be the operator as follows

$$A_nf(x) = \sum_{d(x,y)=n} f(y) = \int_{V_{p+1}} k_n(x,y)f(x)\,\mathrm{d}\mu(x) \text{ with kernel } k_n(x,y) = \begin{cases} 1 & \text{if } d(x,y) = n \\ 0 & \text{otherwise.} \end{cases}$$

These operators have already been introduced in Section 2.3.2 where we also introduced the operators T_t (for $t \in \mathbb{Z}^+$). Proposition 2.28 states for the current situation that

$$T_t = \sum_{0 \leq r \leq t/2} A_{t-2r} = p^{t/2}U_t\left(\frac{A}{2\sqrt{p}}\right).$$

where $U_t(x)$ is the Chebyshev polynomial of the second kind. Let $\mathbb{T}_{p+1}/\Gamma$ be identified with $\mathcal{G}\big(\Lambda(2)/\Lambda(2q)\big)$ and let us restrict the operators A_n to $\Lambda(2q)$ modular functions, i.e. to the finite dimensional space $X^{p,q} \cong \mathbb{T}/\Gamma$. The spectrum of $X^{p,q}$ is nothing but the spectrum of A on $\Lambda(2q)$ modular functions. Then if we write the eigenvalues in the form above the trace formula (Theorem 2.29) claims

$$\sum_{x\in\Lambda(2)/\Lambda(2q)} \sum_{\gamma\Lambda(2q)} k_{t-2r}(x,\gamma x) = p^{t/2}\sum_{j=0}^{n-1} U_t\left(\frac{\mu_j}{2\sqrt{p}}\right) = \sum_{j=0}^{n-1} p^{t/2}\left(\frac{\sin(t+1)\theta_j}{\sin\theta_j}\right).$$

Since $\Lambda(2)\Lambda(2q)$ is vertex-transitive it can be assumed that $x=1$ in the inner sum which yields

$$\sum_{x\in\Lambda(2)/\Lambda(2q)} \sum_{\gamma\Lambda(2q)} k_{t-2r}(x,\gamma x) = |X^{p,q}| \sum_{\gamma\in\Lambda(2q)} k_{t-2r}(1,\gamma) = |X^{p,q}|\Big|\big\{\gamma\in\Lambda(2q)\ \big|\ d(1,\gamma)=t-2r\big\}\Big|.$$

Let $r_Q(\nu)$be the number of representations of ν by Q where Q is a quadratic form defined by

$$Q(x_1,x_2,x_3,x_4) = x_1^2+(2q)^2x_2^2+(2q)^2x_3^2+(2q)^2x_4^2.$$

Then $r_Q(p^k)$ is the number of $\alpha\in H(\mathbb{Z})$ such that $2q|\alpha-a_0$ and $N(\alpha)=p^k$. Corollary 4.8 shows that for every such α is uniquely of the form $\pm p^r R_t(\alpha_1,\overline{\alpha}_1,\dots\alpha_s,\overline{\alpha}_s)$ where $2r+t=k$ and where $[\alpha]\in\Lambda(2q)$. Since the reduced word length in $\Lambda(2)$ corresponds to the distance this observation yields

$$r_Q(p^k) = 2\sum_{r\le k/2}\Big|\big\{\alpha\in\Lambda(2q)\ \big|\ d(1,\alpha)=k-2r\big\}\Big|.$$

Combining the previous results we obtain the remarkable identity

$$r_Q(p^k) = \frac{2p^{k/2}}{|X^{p,q}|}\sum_{j=0}^{n-1}\frac{\sin(k+1)\theta_j}{\sin\theta_j}$$

which also deserves attention in its own right. As stated earlier, the left hand side is the Fourier coefficient of a generalized θ function defined in Section 3.2.1, namely

$$\theta_Q(z) = \sum_{x\in\mathbb{Z}^4} e^{2\pi i Q(x)z} = \sum_{\nu=0}^{\infty} r_Q(\nu)e(\nu z).$$

This is a modular form of weight 2 and hence as a consequence of the structure of $\mathcal{M}_2(\Gamma(16q^2))$ we have the following decomposition

$$\theta_Q(z) = E(z)+f(z)$$

$$E(z) = \sum_{n=0}^{\infty}\delta(n)e(nz)\in\mathcal{E}_2(\Gamma(16q^2)),\ \ f(z)=\sum_{n=1}^{\infty}a(n)e(nz)\in\mathcal{S}_2(\Gamma(16q^2)),$$

which yields the following identity for the Fourier coefficients

$$r_Q(p^k) = \delta(p^k)+a(p^k) = \frac{2p^{k/2}}{n}\sum_{j=0}^{n-1}\frac{\sin(k+1)\theta_j}{\sin\theta_j} = \frac{2p^{k/2}}{n}\sum_{j=0}^{n-1}U_k\left(\frac{\mu_j}{2\sqrt{p}}\right)$$

where $n=|X^{p,q}|$ and

$$\delta(m) = \sum_{d|m} dF(d,m)$$

where F is a suitable periodic function of period $16q^2$ in both variables. We have arrived to the point where the Ramanujan conjecture connects the dots. The term $a(p^k)$ can be estimated by the Ramanujan bound (see Theorem 3.41) as

$$a(p^k) = O\big(p^{k(1/2+\varepsilon)}\big) \text{ as } k \to \infty.$$

Hence we have

$$\delta(p^k) + O\big(p^{k(1/2+\varepsilon)}\big) = \frac{2p^{k/2}}{n} \sum_{j=0}^{n-1} \frac{\sin(k+1)\theta_j}{\sin \theta_j}.$$

To finish the proof the following lemma will be also necessary.

Lemma 4.17. *Let $G : \mathbb{N} \times \mathbb{N} \to \mathbb{C}$ be a periodic function in both variables. Let p be a prime not dividing the period of G. Let $S(k)$ denote the sum $\sum_{d|p^k} dG(d, p^k)$. If $S(k) = o(p^k)$, then $S(k)$ is periodic and hence $S(k) = O_p(1)$.*

Proof. Let us assume that G has period M. Let r be the multiplicative order of p mod M, i.e., r is the smallest positive integer with $p^r \equiv 1 \pmod{M}$. Then for every d, $G(d, p^k) = G(d, p^{k-r})$. If $k \geq r$ we have

$$\frac{S(k) - S(k-r)}{p^k} = \sum_{0 \leq s < r} \frac{G(p^{k-s}, p^k)}{p^s}.$$

By the assumption on $S(k)$,

$$\lim_{k \to \infty} \frac{S(k) - S(k-r)}{p^k} = 0,$$

while the right hand side is periodic in variable k of period r. This implies that $S(k) = S(k-r)$ as it was claimed.

□

Let us assume first that $\left(\frac{p}{q}\right) = -1$. In this case $X^{p,q}$ is bipartite hence by Proposition 2.16 its spectrum is symmetric about zero. $\mu = p+1$ implies $\mu_{n-1} = -(p+1)$ and since $X^{p,q}$ is connected $|\mu_j| < p+1$ for $1 \leq j \leq n-2$. For k odd the right hand side of

$$\delta(p^k) + a(p^k) = \frac{2p^{k/2}}{n} \sum_{j=0}^{n-1} U_k\left(\frac{\mu_j}{2\sqrt{p}}\right)$$

is simply zero, because $U_k(z) = -U(-z)$. If $k = 2$, then the contribution of the trivial eigenvalues can be calculated.

$$U_k\left(\frac{p+1}{2\sqrt{p}}\right) = p^{-\frac{k}{2}} \frac{p^{k+1}-1}{p-1} = p^{-\frac{k}{2}} \sum_{d|p^k} d$$

As we will see, these values provide the leading term of the right hand side.

$$\sum_{d|p^k} dF(d, p^k) + O\big(p^{k(1/2+\varepsilon)}\big) = \frac{4}{n} \sum_{d|p^k} d + \sum_{j=1}^{n-2} U_k\left(\frac{\mu_j}{2\sqrt{p}}\right).$$

On the other hand $U_k(z)$ is a polynomial of degree k, and $|\mu_j| < p+1$ for $j \neq 0, n-1$ hence

$$\sum_{j=1}^{n-2} U_k\left(\frac{\mu_j}{2\sqrt{p}}\right) = o\big(p^{k/2}\big) \text{ as } k \to \infty.$$

This means that for every $\varepsilon > 0$ the following equation holds

$$\sum_{d|p^k} dF(d,p^k) + O\big(p^{k(1/2+\varepsilon)}\big) = \frac{4}{n}\sum_{d|p^k} d + o\big(p^k\big).$$

Rearranging the terms

$$\sum_{d|p^k} d\big(F(d,p^k) - \tfrac{4}{n}\big) = o\big(p^k\big).$$

shows us the conditions of the previous lemma are met. Hence applying Lemma 4.17

$$\delta(p^k) = \begin{cases} O_p(1) & \text{if } k \text{ is odd} \\ \frac{4}{n}\sum_{d|p^k} d + O_p(1) & \text{if } k \text{ is even.} \end{cases}$$

Using this new fact we get a better estimation for the error term of the right hand side by eliminating the leading terms.

$$\frac{2p^{k/2}}{n}\sum_{j=1}^{n-2}\frac{\sin(k+1)\theta_j}{\theta_j} = O\big(p^{k/2+\varepsilon k}\big) \text{ as } k \to \infty.$$

Hence

$$\sum_{j=1}^{n-2}\frac{\sin(k+1)\theta_j}{\theta_j} = O\big(p^{\varepsilon k}\big).$$

Without losing generality we can assume that k is even. Recall $\mu_j = 2\sqrt{p}\cos\theta_j$. Let us assume that some θ_j is not real. Then $|\mu_j| > 2\sqrt{p}$ and we can write

$$\theta_j = \begin{cases} i\phi_j & \text{if } 2\sqrt{p} < \mu_j \le p+1, \\ \pi + i\phi_j & \text{if } -(p+1) \le \mu_j < -2\sqrt{p}, \end{cases}$$

where $0 < \phi_j \le \log\sqrt{p}$. In this case

$$\frac{\sin(k+1)\theta_j}{\theta_j} = \frac{\sin i(k+1)\phi_j}{\sin i\phi_j} = \frac{\sinh(k+1)\phi_j}{\sinh\phi_j} > 0.$$

On the other hand, this quantity cannot be cancelled by the contributions of the real θ's, since we clearly have

$$\left|\frac{2}{n}\sum_{j\,:\,\theta_j\in\mathbb{R}}\frac{\sin(k+1)\theta_j}{\sin\theta_j}\right| \le 2(k+1).$$

Summarizing, if there exists a θ_j which is not real, then choosing a ε small enough and a suitably large even number k, the comparison of the estimates above and the upper bound $O(p^{k\varepsilon})$ leads to a contradiction. This is what had to be demonstrated for $\left(\frac{p}{q}\right) = -1$.

The argument is very similar for $\left(\frac{p}{q}\right) = 1$. In this case $X^{p,q}$ is not bipartite. This time we have

$$\sum_{d|p^k} dF(d,p^k) + O\big(p^{k(1/2+\varepsilon)}\big) = \frac{2(p^{k+1}-1)}{n(p-1)} + o\big(p^k\big)$$

which yields by applying Lemma 4.17 that

$$\sum_{j=1}^{n-1}\frac{\sin(k+1)\theta_j}{\theta_j} = O\big(p^{k\varepsilon}\big) \text{ for all } \varepsilon > 0.$$

□

Since each θ_j is real $\mu(X^{p,q}) \leq 2\sqrt{p}$. This means that the constructed graphs are Ramanujan graphs which was to be demonstrated to finish the proof of Theorem 4.11. $\square$

5. Outlook

5.1 Expanders and geometry

An amazing feature of expander graphs is that they can be viewed from many different angles. Many of the things we have discussed for graphs had been studied earlier in the geometric framework of Riemannian manifolds. In this framework, the continuous analogue of edge expansion can be introduced, and in order to elaborate on the strong connection between the continuous and discrete cases a brief introduction into the spectral theory of the Laplacian operator will be necessary.

5.1.1 Riemannian manifolds in general

A smooth manifold of dimension n is a locally compact, second countable Hausdorff space M equipped with a smooth structure consisting of an open cover $\{U_i\}_{i\in I}$ of M and a collection of homeomorphism $\{\psi_i\}_{i\in I}$ (called local coordinates or charts) such that $\psi_i(U_i)$ is open in $\mathbb{R}^n$ and the charts are smoothly compatible. $T_p(M)$ denotes the tangent space at $p \in M$ and $T_p^*(M)$ is its dual space. TM is the tangent and T^*M is the cotangent bundle. Let $(x^1, \dots x^n)$ be coordinates near $p \in M$ such that $x^k(p) = 0$ for all k. Then the vectors $\{\partial_{x^k}(p)\}_{1\leq k\leq n}$ defined by $\partial_{x^k}(p) = (\delta_k^1, \dots \delta_k^n)$ form a basis of $T_p(M)$. A Riemannian manifold (M, g) is a smooth manifold M with a family of smoothly varying positive definite inner products g_p on $T_p(M)$ for every $p \in M$. The family g is called the Riemannian metric. Since $g_p : T_p(M) \times T_p(M) \to \mathbb{R}$ is a bilinear form, or equivalently it is an element of $T_p^*(M) \otimes T_p^*(M)$ a Riemannian metric is a smooth section of the bundle $T^*(M) \otimes T^*(M)$. The Riemannian metric g is determined by the symmetric, positive definite matrix $(g_{ij}(x)) = (g_p(\partial_{x^i}, \partial_{x^j}))$. We can define the length of a curve $\gamma : [0,1] \to M$ on the Riemannian manifold to be

$$l(\gamma) = \int_0^1 g_p\big(\gamma'(t), \gamma'(t)\big)^{\frac{1}{2}} dt.$$

We define the volume form of an n-dimensional orientable Riemannian manifold to be ω_g which in local coordinates is given by $\omega_g = \sqrt{\det(g_{ij})}\; dx^1 \wedge \dots \wedge dx^n$. The volume of (M, g) is

$$\mu_n(M) = \int_M \sqrt{\det(g_{ij})}\; dx^1 \wedge \dots \wedge dx^n.$$

Let $\Lambda(T_p^*(M))$ denote the exterior algebra of $T_p^*(M)$ and $\Lambda^k(T_p^*(M))$ the homogeneous elements of degree k. $\Lambda^k(T^*M)$ is the exterior k-bundle. A k-form is a smooth section of $\Lambda^k(T^*M)$, i.e., ω is a k-form, if $M \to \Lambda^k(T^*M)$ whose composition with the canonical projection is the identity map. The space of k-forms over M is denoted by $\Omega^k(M)$. We set

$$\Omega(M) = \bigoplus_{0\leq k} \Omega^k(M).$$

Since investigating the differential operators on manifolds is one of our purposes in this section, defining the proper Hilbert space on which the operators will act is still necessary. Let $L^2(M, g)$

be the completion of $C_c^\infty(M)$ with respect to the inner product

$$\langle f, g\rangle = \int_M f(x)\overline{g(x)}\, \mathrm{d}\mu$$

where $C_c^\infty(M)$ is the set of smooth functions of compact support on the manifold M and μ is the volume form defined before. Since T_pM is a finite dimensional vector space with an inner product g_p, it is naturally isomorphic to T_p^*M under the map $\alpha_{g,p} : T_pM \to T_p^*M$, where $\alpha_{g,p}(X_p) = X_p^*$ satisfies $X_p^*(Y_p) = \langle X_p, Y_p\rangle$ for $X_p, Y_p \in T_pM$. We will use α to denote the bundle isomorphism $\alpha : TM \to T^*M$ induced by $\alpha_{g,p}$ in each fiber. In local coordinates, set $g^{ij} = g(dx^i, dx^j)$. In this case $g^{ik}g_{kj} = \delta^i_j$. Furthermore, the inner product g induces canonically an inner product g on each tensor product $T_pM \otimes \cdots \otimes T_pM$ and hence on each exterior power $\Lambda^k T_p^*M$, and hence a global inner product

$$\langle \alpha, \beta\rangle = \int_M g(\alpha, \beta)\, \mathrm{d}\mu$$

for $\alpha, \beta \in C_c^\infty(\Lambda^k T^*M)$. For any $p \in M$ there exists a unique operator d from $\Omega^k(M)$ to $\Omega^{p+1}(M)$, called the exterior derivative and such that for $k = 0$, $d : \mathcal{C}^\infty(M) \to \Omega^1(M)$ is the differential on functions, i.e., $(df)(X) = X(f)$ where X is a vector field on M, for $f \in \mathcal{C}^\infty(M)$, we have $d(df) = 0$ and for $\alpha \in \Omega^p(M)$ and $\beta \in \Omega^q(M)$, we have $d(\alpha\wedge\beta) = d\alpha\wedge\beta + (-1)^p\alpha\wedge d\beta$.

Theorem 5.1 (Stokes theorem for manifolds)**.** *Let M be a compact orientable manifold of dimension n with boundary ∂M and let $\omega \in \Omega^{n-1}(M)$. Then*

$$\int_M d\omega = \int_{\partial M} \omega.$$

The gradient is the operator $\nabla : \mathcal{C}^\infty(M) \to TM$ making $\langle \nabla f, X\rangle = df(X)$ for all $X \in TM$, i.e. ∇ is the composition $\mathcal{C}^\infty(M) \overset{d}{\to} \Lambda^1 T^*M \overset{\alpha^{-1}}{\to} TM$. The gradient vector field is a linear combination of the form $\nabla f = \sum_{k=1}^n c_k\partial_{x^k}$ for some coefficients $c_j \in \mathcal{C}^\infty(M)$. Then

$$\partial_{x^j} f = df(\partial_{x^j}) = \langle \nabla f, \partial_{x^j}\rangle = \Big\langle \sum_{k=1}^n c_k\partial_{x^k}, \partial_{x^j}\Big\rangle = \sum_{k=1}^n a_k g_{kj} = \sum_{k=1}^n \Big(\sum_{i=1}^n g^{ik}\partial_{x^i} f\Big) g_{kj}$$

holds. The uniqueness of the coefficients a_j implies that the gradient operator has the local form

$$\nabla f = \sum_{i,j=1}^n g^{ij}\big(\partial_{x^i} f\big)\partial_{x^j}.$$

Given an n-form $\omega \in \Omega^n(M)$ where $n = \dim(M)$ and any vector field X on M one can define the $(n-1)$-form $\iota_X\omega(X_1, \ldots, X_{n-1}) = \omega(X, X_1, \ldots, X_{n-1})$ where $X_1, \ldots, X_{n-1}$ are arbitrary vector fields on M. Since $d(\iota_X\omega)$ is an n-form there exists a number $\mathrm{div}_\omega X$ such that $d(\iota_X\omega) = (\mathrm{div}_\omega X)\omega$. The divergence operator is the function $\mathrm{div} : TM \to \mathcal{C}^\infty(M)$ such that $d(\iota_x\omega_g) = (\mathrm{div}\, X)\omega_g$. An expression for div in local coordinates can be found similarly to finding one for the gradient operator. For $X \in TM$ we have $X = \sum_{j=1}^n b_j\partial_{x^j}$ with some coefficients $b_j \in \mathcal{C}^\infty(M)$. By definition we have

$$\iota_X\omega_g\Big(\partial_{x^1}\ldots\hat{\partial}_{x^i}\ldots\partial_{x^n}\Big) = \omega_g\Big(X, \partial_{x^1}, \ldots\hat{\partial}_{x^i}, \ldots, \partial_{x^n}\Big)$$

where $\hat{\partial}_{x^i}$ means the only missing term in the expression. Applying the fundamental rules of differential form operations we get

$$(-1)^{i-1}\sqrt{|\det g|}\; dx^1\wedge\cdots\wedge dx^n\Big(\partial_{x^1}, \ldots, X, \ldots\partial_{x^n}\Big) = b_i(-1)^{i-1}\sqrt{|\det g|}.$$

Therefore by the definition of the operator

$$\begin{aligned} d(\iota_X \omega_g) &= d\Big(\sum_{i=1}^n b_i (-1)^{i-1} \sqrt{|\det g|}\ dx^1 \wedge \cdots \wedge \hat{dx^i} \wedge \cdots \wedge dx^n \Big) \\ &= \sum_{i=1}^n (-1)^{i-1} (\partial_{x^i} b_i \sqrt{|\det g|})\ dx^i \wedge dx^1 \wedge \cdots \wedge \hat{dx^i} \wedge \cdots \wedge dx^n \\ &= \sum_{i=1}^n \partial_{x^i} (b_i \sqrt{|\det g|})\ dx^1 \wedge \cdots \wedge dx^n \\ &= \frac{1}{\sqrt{|\det g|}} \sum_{i=1}^n \partial_{x^i} (b_i \sqrt{|\det g|}) \omega_g. \end{aligned}$$

In summary, we see that the operator div has the following local form

$$\operatorname{div} X = \frac{1}{\sqrt{|\det g|}} \sum_{i=1}^n \partial_{x^i} (b_i \sqrt{|\det g|}).$$

The divergence can also be characterized by the equation $\langle -\operatorname{div} X, f\rangle = \langle X, \nabla f\rangle$ for any $f \in \mathcal{C}_c^\infty(M)$ and $X \in TM$.

Theorem 5.2 (Divergence theorem for manifolds)**.** *Let M be an orientable Riemannian manifold and let $X \in TM$. Then*

$$\int_M \operatorname{div} X \omega_g = \int_{\partial M} \langle X, \nu \rangle \sigma_g$$

where ν is the unit vector normal to ∂M and σ_g is the volume form of the boundary.

Proof. By the definition of the div operator, $(\operatorname{div} X)\omega_g = d(\iota_X \omega_g)$. Then by the Stokes theorem

$$\int_M \operatorname{div} X \omega_g = \int_M d(\iota_X \omega_g) = \int_{\partial M} \iota_X \omega_g$$

It is easy to see that $\iota_X \omega_g = \langle X, \nu \rangle \iota_\nu \omega_g = \langle X, \nu \rangle \sigma_{n-1}$ which expresses that the a basis of $T_p \partial M$ can be completed to a basis of $T_p M$ by ν. □

We are now ready to define the Laplace-Beltrami (or Laplacian) operator on functions. The Laplacian operator on (M, g) is the operator $\Delta : \mathcal{C}^\infty(M) \to \mathcal{C}^\infty(M)$ defined as $\Delta = -\operatorname{div} \circ \nabla$. Since both ∇ and div are linear operators it follow for any $f, g \in \mathcal{C}^\infty(M)$ that $\Delta(f+g) = \Delta f + \Delta g$. From the expression of ∇ and div in local coordinates it is straightforward to see that

$$\Delta f = -\frac{1}{\sqrt{|\det g|}} \sum_{i,j=1}^n \partial_{x^i} \Big(g^{ij} \sqrt{|\det g|} \partial_{x^j} f \Big).$$

Note that g^{ij} (and hence g_{ij}) can be recovered by evaluating Δ on a function which is locally $x^i x^j$. This means, that the Laplacian determines the Riemannian metric. This fact suggests that the spectral theory of the Laplace-Beltrami operator is intimately connected to the geometry of the manifold. Before we begin to investigate this connection, we elaborate on the definition of the operator a little.

Theorem 5.3 (Green's theorem for manifolds)**.** *Let (M, g) be a compact orientable Riemannian manifold with boundary ∂M. Then for every $f, h \in \mathcal{C}^\infty(M)$*

$$\int_M f \Delta h \omega_g = \int_M \langle \nabla f, \nabla h \rangle \omega_g - \int_{\partial M} f \nu(h) \sigma_g.$$

Proof. By comparing the local forms of the differential operators div and ∇ we get that for any $\varphi \in \mathcal{C}^\infty(M)$ $\operatorname{div}(\varphi X) = \varphi \operatorname{div} X + \langle \nabla\varphi, X\rangle$. Applying this with the choice $X = \nabla h$

$$\int_M f\Delta h\omega_g = -\int_M \operatorname{div}(f\nabla h)\omega_g + \int_M \langle \nabla f, \nabla h\rangle\omega_g$$

The divergence theorem for manifolds states

$$\int_M \operatorname{div}(f\nabla h)\omega_g = -\int_{\partial M} \langle f\nabla h, \nu\rangle\sigma_g = -\int_{\partial M} f\nu(h)\sigma_g$$

since $dh(\nu) = \nu(h)$. The two equations together prove the claim. □

Corrollary 5.4. *If (M, g) is a compact orientable Riemannian manifold without boundary, then*

(i) $\langle f, \Delta h\rangle = \langle \nabla f, \nabla h\rangle$ *for every* $f, h \in \mathcal{C}^\infty(M)$.

(ii) *the Laplacian operator is formally self-adjoint, i.e.,* $\langle f, \Delta h\rangle = \langle \Delta f, h\rangle$ *for every* $f, h \in \mathcal{C}^\infty(M)$.

(iii) *the Laplacian operator is positive, i.e.,* $\langle \Delta f, f\rangle \geq 0$ *for every* $f \in \mathcal{C}^\infty(M)$.

Let α be the isometry constructed before, in this case $g(\alpha(X), df) = g(X, \nabla f)$. Set $\delta : \Lambda^1 T^*M \mapsto \mathcal{C}^\infty(M)$ by $\delta(\omega) = -\operatorname{div}(\alpha^{-1}(\omega))$. The operation δ is can be characterized by the equation $\langle \delta, \omega\rangle = \langle \omega, df\rangle$ for all $\omega \in \Lambda^1 T^*M$ and $f \in \mathcal{C}_c^\infty(M)$. As a third definition for the Laplacian on functions we have $\Delta = \delta d$. The Hodge star operator, which is a an isometry $* : \Lambda^k T_p^*M \to \Lambda^{n-k} T_p^*M$ such that $*(dx^{i_1} \wedge \cdots \wedge dx^{i_k}) = dx^{j_1} \wedge \cdots \wedge dx^{j_{n-k}}$ where $dx^{i_1} \wedge \cdots \wedge dx^{i_k}$ is a basis element of $\Lambda^k T_p^*M$ and $dx^{i_1} \wedge \cdots \wedge dx^{i_k} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_{n-k}} = \omega_g$. With this notation

$$\langle \alpha, \beta\rangle = \int_M g(\alpha, \beta)\omega_g = \int_M \alpha \wedge *\beta.$$

It also can be shown that for the exterior derivative $d : \Lambda^k T^*M \to \Lambda^{k+1} T^*M$, we have $\langle d\omega, \alpha\rangle = (-1)^{nk+1}\langle \omega, *d * \alpha\rangle$. In particular $\delta = - * d *$. As a result, we are entitled to give a coordinate independent definition for the Laplace-Beltrami operator.

Definition 5.5 (Laplace-Beltrami operator on a manifold)**.** Let (M, g) be a compact orientable Riemannian manifold. Then the Laplacian operator on functions on a Riemannian manifold is given by $\Delta f = \delta df = - * d * df$.

If M is compact and orientable, then the $\operatorname{Ker}(\Delta)$ can be characterized easily. If $\Delta f = 0$ then

$$0 = \int_M \langle \delta df, f\rangle\omega_g = \int_M \langle df, df\rangle\omega_g = \langle df, df\rangle.$$

Therefore, $df = 0$, and f is a constant function. Thus 0 is an eigenvalue of multiplicity one. The solutions of the equation $\Delta f = 0$ are called harmonic functions. For further details one can see [Lab15] and [Bus10].

5.1.2 Spectrum of the Laplace-Beltrami operator

Essentially the same proof presented here can be found in [Ros97]. Assume that (M, g) is a compact Riemannian manifold without boundary. The operator $L = \Delta + \partial_t$ acts on functions in

$\mathcal{C}(M \times (0, +\infty))$ that are $\mathcal{C}^2$ on M and C^1 on $(0, +\infty)$. The heat equation is given by

$$\begin{cases} Lu(x,t) = F(x,t) & (x,t) \in M \times (0,+\infty) \\ u(x,0) = f(x) & x \in M \end{cases}$$

The equation is homogeneous when $F(x,t) \equiv 0$. In this case, the first equation can be reformulated as $\frac{\partial u(x,t)}{\partial t} = -\Delta_x u$. If $u(x,t)$ solves the homogeneous heat equation, then the function $\|u(.,t)\|_{L^2}$ is decreasing with t. Simply because

$$\frac{\partial}{\partial t}\|u(.,t)\|_{L^2}^2 = 2\int_M \partial_t u(x,t)u(x,t)\omega_g(x) = -2\int_M \Delta u(x,t)u(x,t)\omega_g(x) = -2\|\nabla u(.,t)\|^2 \leq 0$$

This implies that the solution to the inhomogenous problem is unique. Suppose that both u_1 and u_2 are solutions, then $u = u_1 - u_2$ solves the homogeneous equation. From the fact that $\int_M u(x,t)^2\omega_g(x)$ is a decreasing function of t and $u(x,0) = 0$ for $x \in M$ follows that $u(x,t) = 0$ for all $x \in M$. The fundamental solution of the heat equation is a continuous function $p : M \times M \times (0,+\infty) \to \mathbb{R}$ which is $\mathcal{C}^2$ with respect to x, $\mathcal{C}^1$ with respect to t and satisfies the conditions $(\Delta_x + \partial_t)p(x,y,t) = 0$ and $\lim_{t\to 0}\int_M p(x,y,t)f(y)\,\mathrm{d}y = f(x)$.

Theorem 5.6 (Existence of heat kernel)**.** *Let M be an n-dimensional compact connected orientable Riemannian manifold without boundary. Then M has a unique fundamental solution $p(x,y,t)$ of the heat equation such that $p(x,y,t) = p(y,x,t)$. The function $p(x,y,t)$ belongs to $\mathcal{C}^\infty(M \times M \times (0,+\infty))$. For $0 < t < 1$, $p(x,y,t)$ obeys the bounds $0 \leq p(x,y,t) \leq c_M t^{-n/2}$, where the constant c_M depends on M. The unique solution $p(x,y,t)$ is called the heat kernel of M.*

The next claim, which is a straightforward consequence of the so-called Duhamel's principle, illustrates the importance of the heat kernel in the theory of heat equations.

Proposition 5.7. *Let $f \in \mathcal{C}(M)$ and $F \in \mathcal{C}(M \times (0,+\infty))$. Then*

$$u(x,t) = \int_M p(x,y,t)f(y)\omega_g(y) + \int_0^t\int_M p(x,y,s)F(y,t-s)\omega_g(y)\mathrm{d}s$$

solves the inhomogenous heat equation

$$\begin{cases} Lu(x,t) = F(x,t) & (x,t) \in M \times (0,+\infty) \\ u(x,0) = f(x) & x \in M \end{cases}$$

Corrollary 5.8 (Properties of the heat kernel)**.** *Let $p(x,y,t)$ be the heat kernel of M.*

$$\int_M p(x,y,t)\omega_g(y) = 1 \text{ and } \int_M p(x,y,t)p(y,z,s)\omega_g(y) = p(x,z,t+s).$$

Proof. The first equation holds because the function $u(x,t) \equiv 1$ solves the equation

$$\begin{cases} Lu(x,t) = 0 & (x,t) \in M \times (0,+\infty) \\ u(x,0) = 1 & x \in M. \end{cases}$$

For fixed $z \in M$ the function $u(x,s) = p(x,z,s+t)$ solves the heat equation with $F \equiv 0$ and $f(y) = p(y,z,t)$ hence

$$p(x,z,t+s) = u(x,s) = \int_M p(x,y,s)p(y,z,t)\omega_g(y).$$

□

At this point we introduce the heat operator $e^{-t\Delta} : L^2(M) \to L^2(M)$ defined by

$$e^{-t\Delta} f(x) = \int_M p(x, y, t) f(y) \omega_g(y)$$

The previous discussion shows that $e^{-t\Delta} f(x)$ solves the homogenous heat equation with initial condition $f(x) \in L^2(M)$. This operator satisfies the identity $e^{-t\Delta} \circ e^{-s\Delta} = e^{-(t+s)\Delta}$. This follows from the fact that $p(x, y, t+s) = \int_M p(x, y, t) p(y, z, s) \omega_g(y)$. The claim has a physical interpretation, the heat flow at time $t+s$ should be the composition of the heat flow up to time t with the heat for a further time s. For $f, g \in L^2(M)$,

$$\begin{aligned} \langle e^{-t\Delta} f, g \rangle &= \int_M \left(\int_M p(x, y, t) f(y) \omega_g(y) \right) g(x) \omega_g(x) \\ &= \int_M \left(\int_M p(y, x, t) g(x) \omega_g(x) \right) f(y) \omega_g(y) = \langle f, e^{-t\Delta} g \rangle \end{aligned}$$

i.e., the operator is self-adjoint. Since $\langle e^{-t\Delta} f, f \rangle = \|e^{-\frac{t}{2}\Delta} f\|^2 \geq 0$ the operator is also positive. Furthermore, it is a compact operator as a consequence of the Rellich-Kondarachov compactness theorem about the inclusion between the Sobolev spaces of Riemannian manifolds. As $t \to 0$ we have $e^{-t\Delta} f \to f$ in $L^2(M)$. If t is sufficiently small, then

$$\left\| f - \int_M p(x, y, t) f(y) \omega_g(y) \right\|_{L^2}^2 \leq \int_M \int_M p(x, y, t) |f(x) - f(y)|^2 \omega_g(y) \omega_g(x).$$

The functions $p(x, y, t)|f(x) - f(y)|^2$ are uniformly bounded and tend to 0 as $t \to 0$. The claimed statement can be concluded by the application of the dominated convergence theorem.

Lemma 5.9. *Let M be a compact connected orientable Riemannian manifold without boundary.For any $f \in L^2(M)$ the function $e^{-t\Delta} f$ converges uniformly as $t \to \infty$ to a constant function.*

Proof. As we have already shown, $\|e^{-t\Delta} f\|_{L^2}$ is a decreasing function of t so it is convergent in $L^2(M)$. Let $x \in M$, we have

$$\left| \left(e^{-(t+T)\Delta} f - e^{-(s+T)\Delta} f \right)(x) \right|^2 = \left| \int_M p(x, y, T) \left(e^{-t\Delta} f - e^{-s\Delta} f \right)(y) \omega_g(y) \right|^2$$

There exists a suitable constant $C_1 > 0$ such that the right-hand side is at most $C_1 \|e^{-t\Delta} f - e^{-S\Delta} f\|_{L^2}^2$. Here $\|e^{-t\Delta} f - e^{-S\Delta} f\|_{L^2}^2 \to 0$ as $s, t \to \infty$, which can be seen by

$$\|e^{-t\Delta} f - e^{-s\Delta} f\|_{L^2}^2 = \|e^{-t\Delta} f\|_{L^2}^2 - 2 \left\| e^{-\frac{t+s}{2}\Delta} f \right\|_{L^2}^2 + \|e^{-t\Delta}\|_{L^2}^2$$

and the fact that $e^{-t\Delta} f$ is convergent in $L^2(M)$. The above upper bound implies that $e^{-t\Delta} f$ also converges uniformly necessarily to a continuous function φ, i.e., $\lim_{t \to \infty} e^{-t\Delta} f = \varphi$. Now we show that $e^{-\Delta} \varphi = \varphi$.

$$\lim_{t \to \infty} \left| \left(e^{-(t+s)\Delta} f - e^{-t\Delta} \varphi \right)(x) \right|^2 = \lim_{t \to \infty} \left| \int_M p(x, y, t) \left(e^{-s\Delta} f - \varphi \right)(y) \omega_g(y) \right|^2 = 0$$

since the square in the last expression can be bounded by $C_2 \|e^{-s\Delta} f - \varphi\|_{L^2}^2$ with a suitable $C_2 > 0$ where the L^2 norm tends to 0. By the definition of the heat operator $\varphi(x, t) = e^{-t\Delta} f(x)$ solves the heat equation with the initial condition $\varphi(0, x) = \varphi$. Considering the fact that $e^{-t\Delta} \varphi = \varphi$ one can conclude that $(\Delta_x + \partial_t) \varphi(x) = 0$ which means that $\Delta_x \varphi(x) = 0$, i.e., the functions is harmonic on M. Since M is compact, the constant functions are the only harmonic functions on the manifold. □

Theorem 5.10 (Spectrum of the Laplacian)**.** *Let (M, g) be a compact connected orientable Riemannian manifold without boundary. Let Δ be the Laplace-Beltrami operator on the manifold. The eigenvalue problem $\Delta\phi = \lambda\phi$ $(\phi \in \mathcal{C}^\infty(M), \lambda \in \mathbb{R})$ has a complete orthonormal system of smooth eigenfunctions $\{\phi_j\}_{0\leq j}$ in $L^2(M)$ with corresponding eigenvalues $\{\lambda_j\}_{0\leq j}$. These have following properties*

(i) $0 = \lambda_0 \leq \lambda_1 \leq \dots$ *and* $\lim_{n\to\infty} \lambda_n = +\infty$.

(ii) $p(x, y, t) = \sum_{n=0}^{\infty} e^{-\lambda_n t}\phi_n(x)\phi_n(y)$ *where the series converges uniformly on $M \times M$ for $t > 0$.*

Proof. The first step of the proof is to show $e^{-t\Delta} = (e^{-\Delta})^t$ for $t \in (0, +\infty)$. If $k \in \mathbb{Z}_+$ then this property follows from the identity $e^{-s\Delta} \circ e^{-t\Delta} = e^{-(t+s)\Delta}$. It follows similarly that for $e^{-\frac{1}{k}\Delta} = \left(e^{-\Delta}\right)^{\frac{1}{k}}$.

$$\left\|e^{-t\Delta} - \left(e^{-\Delta}\right)^t\right\| \leq \left\|e^{-t\Delta} - e^{-\frac{p}{q}\Delta}\right\| + \left\|e^{-\frac{p}{q}\Delta} - \left(e^{-\Delta}\right)^{\frac{p}{q}}\right\| + \left\|\left(e^{-\Delta}\right)^{\frac{p}{q}} - \left(e^{-\Delta}\right)^t\right\|$$

The second term is obviously 0. The following estimation can be given to the first term

$$\left\|\left(e^{-t\Delta} - e^{-\frac{p}{q}\Delta}\right)f\right\|_{L^2}^2 \leq \int_M \int_M \left|p(x, y, t) - p(x, y, p/q)\right|^2 f(y)\omega_g(y)\omega_g(x)$$

which converges to 0 by the Dominated Convergence Theorem. As for the last term

$$\lim_{p/q\to t} \left\|\left(e^{-\Delta}\right)^{\frac{p}{q}} - \left(e^{-\Delta}\right)^t\right\| = \lim_{p/q\to t} \left(\sup_{\beta\in \mathrm{Sp}(e^{-\Delta})} \left|\beta^{\frac{p}{q}} - \beta^t\right|\right) = 0.$$

By the spectral theorem for self-adjoint compact operators on Hilbert spaces, $L^2(M)$ has an orthonormal basis consisting of eigenfunctions $\{\phi_j\}_{0\leq j}$ for the heat operator $e^{-\Delta}$ with eigenvalues $\beta_j \to 0$ as $j \to \infty$. We certainly have $\beta_j \geq 0$ but we show $\beta_j > 0$ also holds. Assume by contradiction that for some $j \geq 0$ $\beta_j = 0$, then there exists $f \neq 0$ with $e^{-\Delta}f = 0$.

$$0 = \left\langle e^{-t\Delta}f, f\right\rangle = \left\langle e^{-\frac{1}{n}\Delta}f, f\right\rangle$$

which implies $e^{-\frac{1}{n}\Delta}f = 0$ and hence $f = \lim_{t\to 0} e^{-t\Delta}f = 0$, a contradiction. Therefore $\beta_j > 0$. Now set $\lambda_j := -\log\beta_j$. Then

$$e^{-t\Delta}\phi_j = e^{-t\lambda_j}\phi_j.$$

Since $e^{-t\Delta}\phi_j$ is a solution of the heat equation for all j, we get

$$0 = (\Delta + \partial_t)\left(e^{-t\Delta}\phi_j\right) = e^{-\lambda_j t}(\Delta\phi_j + \lambda_j\phi_j)$$

which implies that ϕ_j is an eigenfunction of the Laplace-Beltrami operator with eigenvalue λ_j. In order to prove **(ii)** note that

$$p(x, y, t) = \sum_{j=0}^{\infty}\langle p(x, ., t), \phi_j\rangle\phi_j(y) = \sum_{j=0}^{\infty} e^{-t\Delta}\phi_j(x)\phi_j(y) = \sum_{j=0}^{\infty} e^{-\lambda_j t}\phi_j(x)\phi_j(y).$$

□

The eigenvalues of the Laplace-Beltrami operator on the manifold will be denoted by $\lambda_j(M)$ for $0 \leq j$. Using this convention, $\lambda_1(M)$ denotes the smallest positive eigenvalue of the Laplacian.

Theorem 5.11 (Rayleigh's principle)**.** *Let (M, g) be a compact connected Riemannian manifold. For each $1 \leq j$ we have*

$$\lambda_j(M) = \inf\left\{\frac{\int_M \|df\|^2\omega_g}{\int_M |f|^2\omega_g} \;\Big|\; f \in \mathcal{C}^\infty(M) \text{ and } \langle f, \phi_0\rangle = \cdots = \langle f, \phi_{j-1}\rangle = 0\right\}.$$

Theorem 5.12 (Weyl formula)**.** *Let (M, g) be a compact connected Riemannian manifold. Let $\{\lambda_j\}_{j=0}^\infty$ be the eigenvalues with multiplicity of the Laplace-Beltrami operator Δ. Then*

$$\sum_{\lambda_k \leq \lambda} 1 \sim \frac{\omega_n \mu_n(M)}{(2\pi)^n}\lambda^{\frac{n}{2}} \text{ and } \lambda_k \sim \frac{(2\pi)^2}{\omega_n^{2/n}}\Big(\frac{k}{\mu_n(M)}\Big)^{\frac{2}{n}} \text{ as } k \to \infty$$

where $\mu_n(M)$ denotes the Riemannian volume of M and ω_n is the n-dimensional volume of the unit ball in $\mathbb{R}^n$.

Theorem 5.13 (Sup-norm bound)**.** *Let (M, g) be a compact Riemannian manifold of dimension n. Then if $\Delta\phi = \lambda\phi$,*

$$\|\phi\|_{L^\infty} \ll_M \lambda^{\frac{n-1}{4}}.$$

5.1.3 Cheeger inequality

Let $\lambda_1(M)$ be the smallest positive Laplacian eigenvalue. We have the following characterization of $\lambda_1(M)$ independent of Δ.

Proposition 5.14. *Let M be a manifold without boundary. Then*

$$\lambda_1(M) = \inf\left\{\frac{\int_M \|df\|^2}{\int_M |f|^2} \;\Big|\; f \in \mathcal{C}^\infty(M) \text{ and } \int_M f = 0\right\}$$

Definition 5.15 (Cheeger constant)**.** Let (M, g) be an n-dimensional compact Riemannian manifold. For $l \in \mathbb{N}^+$ let μ_l be the l-dimensional volume of a submanifold. The Cheeger constant of M is

$$h(M) = \inf_{N \subset M} \frac{\mu_{n-1}(\partial N)}{\min\big(\mu_n(N), \mu_n(M \setminus N)\big)}$$

where the infimum is taken over all compact n-dimensional submanifolds N whose boundary ∂N is an $(n-1)$-dimensional submanifold and $M \setminus N$ is a submanifold of dimension n.

The analogy with the definition of edge expansion is obvious. Furthermore, this is the generalization of the classical isoperimetric problem on the plane.

Theorem 5.16 (Cheeger inequality)**.** *Let M be a compact orientable Riemannian manifold, and let $\lambda_1(M)$ be the smallest positive eigenvalue of its Laplacian operator. Then $\lambda_1(M) \geq \frac{h(M)^2}{4}$.*

The proof of this theorem can be found in [Bus10] or in [Led94]. On the other hand, an upper bound for $\lambda_1(M)$ can be also established. Let $R(M)$ be the Ricci curvature.

Theorem 5.17 (Buser)**.** *If $R(M) \geq -(n-1)a^2$ for some $a \geq 0$ where $n = \dim M$, then $\lambda_1(M) \leq 2a(n-1)h(M) + 10h^2(M)$.*

In analogy with the geometric Laplacian, one can define a discrete Laplacian operator for graphs. Let $G = (V, E)$ be a finite graph. Let us assume that there is a fixed but arbitrary orientation on the edges of the graph. For $e \in E$ let e_- its origin and e_+ its target. The operator $d : L^2(V, \mu) \to L^2(E, \mu)$ is defined as $df(e) = f(e_+) - f(e_-)$. If $|V| = n$ and $|E| = m$ then the matrix of d is an $m \times n$ matrix D indexed by pairs $(e, v) \in E \times V$ such that

$$D_{e,v} = \begin{cases} 1 & \text{if } v = e_+ \\ -1 & \text{if } v = e_- \\ 0 & \textit{otherwise.} \end{cases}$$

Definition 5.18 (Discrete Laplacian operator)**.** The discrete Laplacian operator $\Delta : L^2(G, \mu) \to L^2(G, \mu)$ is defined by the matrix $\Delta = D^T D$ where D^T is the transpose of D.

Let S be the $n \times n$ diagonal matrix indexed by $V \times V$ where $S_{v,v} = \deg(v)$. Then $\Delta = S - A$, where A is the adjacency matrix of G. It is not difficult to show that Δ is positive, so its eigenvalues are non-negative. If G is finite, then zero is always an eigenvalue with the constant function as an eigenfunction. The eigenfunctions of the Laplacian operator are called harmonic functions. In the light of this fact, the only harmonic functions are the constants on a graph. The zero is a simple eigenvalue if and only if G is connected. Let $\lambda_1(G)$ be the smallest positive eigenvalue. In analogy with Cheeger's inequality we have the following lower bound on $\lambda_1(G)$.

Proposition 5.19. *Let G be a finite graph with* $\deg(G) \leq d$. *Then* $\lambda_1(G) \geq \frac{h^2(G)}{2d}$.

In analogy with Buser's theorem, one can derive a lower bound on $\lambda_1(G)$ also. See in [HLW06] and in [Lub10].

Proposition 5.20. *Let G be a finite graph with* $\deg(G) \leq d$. *Then* $h(G) \geq \lambda_1(G)/2$.

Corrollary 5.21 (Spectral characterization of expander families by the Laplacian operator)**.** *A family $(G_i)_{i \in I}$ of non-empty connected finite graphs $G_i = (V_i, E_i)$ is an expander family, if there exists constant $\varepsilon > 0$, such that:*

(i) *For any $n \geq 1$, there are only finitely many $i \in I$ such that $|V_i| \leq n$.*
(ii) $\sup_{i \in I} \deg(G_i) < \infty$.
(iii) *For each $i \in I$, the spectral gap satisfies $\lambda_1(G_i) \geq \varepsilon$.*

5.2 Expanders and representation theory

In this last chapter a new problem will be presented, namely the Banach-Ruziewich problem. It asks whether the Lebesgue measure is the only finitely additive measure of total measure one, defined on the Lebesgue measurable subsets of the n-dimensional sphere and invariant under rotations. At first glance, this new topic is completely unrelated to the problem of expander graphs, but in fact, both problems were solved by similar methods, initially, by Kazhdan's property (T), and later by using the Ramanujan conjecture from the theory of automorphic forms. The Kazhdan's property (T) will be the central concept of this part.

5.2.1 Kazhdan's property (T)

A (linear) representation of a group Γ in a vector space V is a homomorphism from Γ to the group of linear automorphisms of V. Let $\mathcal{U}(H)$ be the unitary group of H, i.e., the group of all invertible unitary operators $U : H \to H$. A unitary representation of a topological group Γ in a Hilbert space H is a pair (π, H) where $\pi : \Gamma \to \mathcal{U}(H)$ is a group homomorphism such that $\gamma \to \pi(\gamma)v$ is continuous for every vector $v \in H$. An invariant subspace of a representation (π, H) is a vector subspace $U \subseteq H$ such that $\pi(\gamma)U \subseteq U$ for all $\gamma \in \Gamma$. A representation is said to be irreducible if it has no closed invariant subspace other than 0 and H. Let (π, H) be a unitary representation of Γ and let $U \subseteq H$ be a closed Γ-invariant subspace. For every $\gamma \in \Gamma$, denoting by $\pi^U(\gamma) : U \to U$ the restriction of the operator $\pi(\gamma)$ to U, we obtain a unitary representation π^U of Γ on U. In this case, (π^U, U) is a subrepresentation of (π, H). The orthogonal complement $U^\perp$ of U is also Γ-invariant since $\langle \pi(\gamma)v, w\rangle = \langle w, \pi^*(\gamma)w\rangle = \langle v, \pi(\gamma^{-1})w\rangle = 0$ holds for every $\gamma \in \Gamma$, $v \in U^\perp$ and $w \in U$. Let π_0 denote the one-dimensional trivial representation.

Definition 5.22 (Almost invariant vectors)**.** Let (π, H) be a unitary representation of a topological group Γ. For a subset Q of Γ and a real number $\varepsilon > 0$ a vector $w \in H$ is (Q, ε)-invariant if

$$\sup_{\gamma \in Q} \|\pi(\gamma)w - w\| < \varepsilon \|w\|.$$

The representation (π, H) has almost invariant vectors if it has (Q, ε)-invariant vectors for every compact subset Q of Γ and every $\varepsilon > 0$. If this holds the trivial representation is said to be weakly contained in π. This is denoted by $\pi_0 \prec \pi$.

Definition 5.23 (Invariant vectors)**.** The representation (π, H) has non-zero invariant vector if there exists $w \neq 0$ in H such that $\pi(\gamma)w = w$ for all $\gamma \in \Gamma$. If this holds, we write $\pi_0 \subset \pi$. This means that π_0 is a subrepresentation of π.

Definition 5.24 (Kazhdan's property (T))**.** Let Γ be a topological group. A subset $Q \subset \Gamma$ is said to be a Kazhdan set if there exists $\varepsilon > 0$ with the following property: every unitary representation (π, H) of Γ which has a (Q, ε)-invariant vector has a non-zero invariant vector. In this case $\varepsilon > 0$ is called a Kazhdan constant for Γ and Q, and (Q, ε) is called a Kazhdan pair for Γ. The group Γ has Kazhdan's property (T) if Γ has a compact Kazhdan set. Or equivalently Γ is said to be a Kazhdan group.

The (T) in the name refers to the trivial representation since the unitary dual, i.e., the set of equivalence classes of irreducible unitary representations, can be equipped with a topology, the so-called Fell topology such that the group is Kazhdan if and only if the unitary representation is an isolated point in the unitary dual. It is not completely trivial to exhibit a Kazhdan group but for example compact groups have Kazhdan's property (T). Since free groups do not have this property, it can be shown by applying the ping-pong lemma that $SL_2(\mathbb{R})$ does not have property (T).

Theorem 5.25 (Kazhdan)**.** *If Λ is a lattice in Γ, then Γ has property (T) if and only if Λ has.*

5.2.2 The Banach-Ruziewicz problem

Let us assume that a group Γ acts on a set X, and $A, B \subset X$. A and B are said to be γ-equidecomposable if there exist pairwise disjoint sets $\{A_j\}_{j=1}^k$ of A and pairwise disjoint sets $\{B_j\}_{j=1}^k$ of B such that $A = \cup_{j=1}^k A_j$, $B = \cup_{j=1}^k B_j$ and there are $\gamma_1, \ldots, \gamma_k \in \Gamma$ such that for

$1 \le j \le k$, $\gamma_j(A_j) = B_j$, i.e., A and B can each be partitioned into the same finite number of respectively Γ-congruent pieces. This fact will be denoted by $A \cong_\Gamma B$. The set X said to be Γ-paradoxical if there are two proper disjoint subsets $A, B \subset X$ such that $A \cong_\Gamma X \cong_\Gamma B$. Or equivalently $A \cup B = X$ and $A \cong_\Gamma X \cong_\Gamma B$. This means that X can be decomposed into finitely many pieces from which two copies of X can be rebuilt. Let $F = \langle a_1, a_2 \rangle$ be a free group of rank 2. F acts on itself by left multiplication. For $j = 1, 2$, let A_j^ε be the set of all reduced words beginning with a_j^ε, where $\varepsilon = \pm 1$. Then clearly $A_j^{+1} \cup a_j A_j^{-1} = F$ hence $F \cong_F \left(A_j^{+1} \cup A_j^{-1}\right)$ for $j = 1, 2$ which means that $F \cong_F F$. In an arbitrary group Γ, a subset $S \subseteq \Gamma$ is called independent if S is a free set of generators of the subgroup $\langle S \rangle$. It is a well-known fact, that there exist independent elements in $\mathrm{Aut}(S^2)$. For example the rotations described by the matrices

$$A_1 = \begin{pmatrix} \frac{1}{3} & \frac{2\sqrt{2}}{3} & \\ -\frac{2\sqrt{2}}{3} & \frac{1}{3} & \\ & & 1 \end{pmatrix} \text{ and } A_2 = \begin{pmatrix} 1 & & \\ & \frac{1}{3} & \frac{2\sqrt{2}}{3} \\ & -\frac{2\sqrt{2}}{3} & \frac{1}{3} \end{pmatrix}$$

are independent. As a consequence, the group $SO_3(\mathbb{R})$ contains a free group on two generators.

Proposition 5.26. *S^2 is $SO_3(\mathbb{R})$-paradoxical.*

Proof. The steps of the proof are the following: **(i)** firstly we show that $S^2 \cong_{SO_3(\mathbb{R})} (S^2 \setminus D)$, where D is a countable set, **(ii)** then it will be proven that $S^2 \setminus D$ is $SO_3(\mathbb{R})$-paradoxical. If $(S^2 \setminus D)$ is $SO_3(\mathbb{R})$-paradoxical, then by definition there are two proper disjoint subsets A, B of $(S^2 \setminus D)$ such that $A \cong_{SO_3(\mathbb{R})} (S^2 \setminus D) \cong_{SO_3(\mathbb{R})} B$. Since $(S^2 \setminus D) \cong_{SO_3(\mathbb{R})} S^2$, the relation $A \cong_{SO_3(\mathbb{R})} S^2 \cong_{SO_3(\mathbb{R})} B$ also holds, which means that S^2 is $SO_3(\mathbb{R})$-paradoxical as well.

(i) Let $\langle a_1, a_2 \rangle = F \le SO_3(\mathbb{R})$ be a free subgroup of rank 2. Let $D = \{x \in S^2 \mid \exists 1 \ne \gamma \in F, \gamma(x) = x\}$. This set is countable since F is countable and every non-trivial element of $SO_3(\mathbb{R})$ has exactly 2 fixed points in S^2. Let ℓ be a line through the origin that misses the countable set D. Let $R_\vartheta \in SO_3(\mathbb{R})$ be a rotation about ℓ by angle ϑ. The set $\mathcal{R} = \{\vartheta \in \mathbb{R} \mid \exists n > 0, R_\vartheta^n(D) \cap D \ne \emptyset\}$ is clearly countable, hence there exists $\vartheta \notin \mathcal{R}$, i.e., an angle such that $R_\vartheta^n(D) \cap D$ is empty for all $n > 0$. From this fact, it follows that the sets $D, R_\vartheta(D), R_\vartheta^2(D) \ldots$ are pairwise disjoint. Taking $D' = \bigcup_{j=0}^n R_\vartheta^n(D)$ we proved our first claim since

$$S^2 = D' \cup (S^2 \setminus D') \cong_{SO_3(\mathbb{R})} R_\vartheta(D') \cup (S^2 \setminus D') = S^2 \setminus D.$$

(ii) F acts freely on $S^2 \setminus D$, i.e., no $\gamma \in F$ has a fixed point. Let $\mathcal{M}$ be a set of representatives for the F-orbits in $S^2 \setminus D$. Let A_j^ε be as before, the set of reduced words beginning with a_j^ε where $\varepsilon = \pm 1$. The sets $A_j^\varepsilon(\mathcal{M})$ are disjoint since F acts freely and

$$(S^2 \setminus D) = A_1^{+1}(\mathcal{M}) \cup a_1 A_1^{-1}(\mathcal{M}) = A_2^{+1}(\mathcal{M}) \cup a_2 A_2^{-1}(\mathcal{M})$$

which implies that $S^2 \setminus D$ is F-paradoxical and hence $SO_3(\mathbb{R})$-paradoxical. □

By induction it can be proved that S^n is $SO_{n+1}(\mathbb{R})$-paradoxical for every $n \ge 2$. In addition, if $n \ge 2$ then any two subsets of S^n, each of which has a non-empty interior, are $SO_{n+1}(\mathbb{R})$-equidecomposable. In particular S^n is equidecomposable with each of its subsets with a non-empty interior.

Theorem 5.27 (Tarski)**.** *Assume that* Γ *acts on* X*. Then there is a finitely additive* Γ*-invariant measure* $\nu : P(X) \to [0,1]$ *with* $\nu(X) = 1$ *if and only if* X *is not* Γ*-paradoxical.*

Proposition 5.28. *Let* ν *be an* $SO_{n+1}(\mathbb{R})$*-invariant measure on Lebesgue measurable subset of* S^n *such that* $\nu(S^n) = 1$*. Then* $\nu = \lambda$*, where* λ *is the normalised Lebesgue measure.*

This claim follows from the uniqueness of the Haar measure. Let $n \geq 1$. The Banach-Ruziewich problem asks whether the normalised Lebesgue measure, defined on all Lebesgue measurable subsets of S^n is the unique normalised $SO_{n+1}(\mathbb{R})$-invariant finitely additive measure. In the case $n = 1$, it was shown by Banach that the answer is negative.

Definition 5.29 (Invariant mean)**.** Let Γ be a locally compact group and $L^\infty(\Gamma)$ the equivalence classes of real-valued measurable functions bounded outside a set of zero Haar measure. An invariant mean on Γ is a linear functional $m : L^\infty(\Gamma) \to \mathbb{R}$ satisfying:

(i) $m(f) \geq 0$ if $f \geq 0$.

(ii) $m(\chi_\Gamma) = 1$ where χ_Γ is the characteristic function of Γ.

(iii) $m({}_g f) = m(f)$ for every $g \in \Gamma$ and $f \in L^\infty(\Gamma)$ where $({}_g f)(x) = f(gx)$.

A mean is automatically continuous since $|m(f)| \leq \|f\|_{L^\infty}$ If there exists an invariant mean m on a group Γ, then by defining, for a subset $A \subseteq \Gamma$ a measure $\nu(A) = m(\chi_A)$, we obtain a finitely additive Γ-invariant measure defined on all subsets of Γ. A groups is amenable if there exists an invariant mean on it. According to Følner's theorem, the amenability of a group can be also characterized by the property (F), namely: given $\varepsilon > 0$ and a compact set $K \subset G$ there is a Borel set $U \subseteq \Gamma$ of positive finite left Haar measure $\mu(U)$ such that $\frac{1}{\mu(U)}\mu(gU\Delta U) < \varepsilon$ for all $g \in K$. To establish a connection with the previous material, the amenability of a discrete group Γ means that for every $\varepsilon > 0$ and every compact set $K \subset \Gamma$, the Cayley graph $\mathcal{G}(\Gamma, K)$ has a finite subset U of vertices whose boundary ∂U satisfies $|\partial U| < \varepsilon |U|$. The condition of amenability is somewhat the opposite of the expander property.

Theorem 5.30 (Banach)**.** *Let* X *be a dense* G_δ *subset of* S^1 *such that* $\lambda(X) = 0$*. Then there exists a finitely additive invariant measure* ν *defined on all subsets of* S^1 *such that* $\nu(X) = 1$*.*

On the other hand, Tarski showed, using the Banach-Tarski paradox, that the situation for $n \geq 2$ is more rigid.

Theorem 5.31 (Tarski)**.** *Let* ν *be an* $SO_{n+1}(\mathbb{R})$*-invariant finitely additive measure on the Lebesgue measurable subsets of* S^n*,* $n \geq 2$*. Then* ν *is absolutely continuous with respect to* λ*.*

Let $\pi_\nu : \Gamma \to L^2(\Omega, \nu)$ be a unitary representation defined by $\pi_\nu(g)f(\omega) = f(g^{-1}\omega)$ where (Ω, ν) is a probability measure space and $\omega \in \Omega$. Since $\chi_\Omega \in L^2(\Omega, \nu)$ we have

$$L^2(\Omega, \nu) = \mathbb{R}\chi_\Omega \bigoplus \left\{ f \in L^2(\Omega, \nu) \;\middle|\; \int_\Omega f(\omega)\, \mathrm{d}\nu(\omega) = 0 \right\}.$$

The orthogonal component of $\mathbb{R}\chi_\Omega$ is denoted by $L_0^2(\Omega, \nu)$ which is Γ-invariant and $\pi_\nu^0 = \pi_\nu|_{L_0^2(\Omega,\nu)}$. An action of Γ on Ω is ergodic if $\nu(A) = 0$ or $\nu(\Omega \setminus A) = 0$ for any Γ measurable subset $A \subseteq \Omega$. We will use the fact that the ergodicity of an action is equivalent to the fact that π_ν^0 has no non-zero invariant functions. The following result, due to Rosenblatt, Schmidt and Losert-Rindler relates the Banach-Ruziewicz problem to property (T) and plays a key role in its solution.

Proposition 5.32 (Banach-Ruziewich problem and unitary representations)**.** *Let* Γ *be a countable group acting in a measure preserving way on a probability space* (Ω, ν)*. If the associated unitary representation* π_ν^0 *does not contain the trivial representation weakly then* $\nu(f) = \displaystyle\int_\Omega f(g)\ \mathrm{d}\nu(g)$ *is the unique* Γ*-invariant mean on* $L^\infty(\Omega, \nu)$*.*

Proof. Let m be a Γ-invariant mean on $L^\infty(\Omega, \nu)$. We are supposed to show that $m = \nu$. Let $\mathcal{M} \subseteq L^\infty(\Omega, \nu)^*$ be the set of all means on $L^\infty(\Omega, \nu)$. $\mathcal{M}$ is a closed subset of the unit ball in the dual space with respect to the weak*-topology. According to the Banach-Alaoglu theorem, the unit ball of the dual space is compact, hence $\mathcal{M}$ is also compact. Let

$$L^1(\Omega, \nu)_{1,+} = \left\{ f \in L^1(\Omega, \nu) \;\middle|\; f(\omega) \geq 0 \text{ for all } \omega \in \Omega \text{ and } \int_\Omega f(\omega)\ \mathrm{d}\nu(\omega) = 1 \right\}.$$

If $\varphi \in L^1(\Omega, \nu)_{1,+}$ then

$$\int_\Omega f(\omega) \mathrm{d}\varphi(\omega) \in \mathcal{M}$$

and hence $L^1(G)_{1,+}$ can be viewed as a subset of $\mathcal{M}$. As a consequence of the Hahn-Banach theorem, it is a weak* dense subset in $\mathcal{M}$. Hence a net $(f_i)_{i \in I}$ can be found in $L^1(\Omega, \nu)_{1,+}$ converging to m in the weak*-topology. Since m is Γ-invariant, we have $\left({}_g f_i - f_i\right) \xrightarrow[i \in I]{w^*} 0$ for all $g \in \Gamma$ in the weak topology of $L^1(\Omega, \nu)$. Let us consider the product space

$$\mathcal{X} = \prod_{g \in \Gamma} L^1(\Omega, \nu)$$

with the product of the norm topologies. Since the projections $\{p_g\}_{g \in \Gamma}$ generate the topology and satisfy

$$\bigcap_{g \in \Gamma} \{x \in \mathcal{X} \mid p_f(x) = 0\} = \{0\},$$

$\mathcal{X}$ is locally convex. In addition the weak topology on it is the product of the weak topologies on the factors. The set

$$\mathcal{A} = \left\{ ({}_g f - f)_{g \in \Gamma} \mid f \in L^1(\Omega, \nu)_{1,+} \right\} \subset \mathcal{X}$$

is convex and its closure in the weak topology contains 0. The weak closure of a convex subset of a locally convex space is equal to its original closure. Hence, there exists a net $(f_i')_{i \in I}$ in $L^1(\Omega, \nu)_{1,+}$ such that for every $g \in \Gamma$ $\left({}_g f_i' - f_i'\right) \xrightarrow[i \in I]{L^1} 0$. By definition of $L^1(\Omega, \nu)_{1,+}$ it is possible to consider the functions $\varphi_i = \sqrt{f_i'}$ which are clearly element of $L^2(\Omega, \nu)$ with $\|\varphi_i\|_{L^2(\Omega, \nu)}$.

$$\begin{aligned} \|\pi_\nu(g)\varphi_i - \varphi_i\|_{L^2(\Omega,\nu)} &= \int_\Omega \left|\varphi_i(g^{-1}\omega) - \varphi_i(\omega)\right|^2 \mathrm{d}\nu(\omega) \\ &\leq \int_\Omega \left|\varphi_i(g^{-1}\omega)^2 - \varphi_i(\omega)^2\right| \mathrm{d}\nu(\omega) = \|{}_{g^{-1}} f_i' - f_i'\|_{L^1(\Omega,\nu)} \end{aligned}$$

This upper bound shows that $\left(\pi_\nu(g)\varphi_i - \varphi_i\right) \xrightarrow[i \in I]{L^2} 0$ for all $g \in \Gamma$. Let $\varphi_i = c_i \chi_\Omega + \psi_i$ be the orthogonal decomposition of φ_i where $\psi_i \in L_0^2(\Omega, \nu)$ and $c_i = \int_\Omega \varphi_i(\omega)\ \mathrm{d}\nu(\omega)$. We have

$\left(\pi_\nu^0(g)\psi_i - \psi_i\right) \xrightarrow[i\in I]{L^2} 0$ for all $g \in \Gamma$. By the assumptions of the proposition, the trivial representation is not weakly contained in π_ν^0 which implies that $\inf_{i\in I} \|\psi_i\|_{L^2(\Omega,\nu)} = 0$. We can assume, upon choosing a proper subnet, that $\psi_i \xrightarrow[i\in I]{L^2} 0$. It follows that $\lim_{i\in I} c_i = 1$ and hence $\varphi_i \xrightarrow[i\in I]{L^2} \chi_\Omega$. Then by

$$\|f_i' - \chi_\Omega\|_{L^1(\Omega,\nu)} = \int_\Omega |\varphi_i(\omega) - 1||\varphi_i(\omega) + 1| \,\mathrm{d}\nu(\omega) \leq 2\|\varphi_i - \chi_\Omega\|_{L^2(\Omega,\nu)}$$

we have $f_i' \xrightarrow[i\in I]{L^2} \chi_\Omega$. By comparing this with $f_i' \xrightarrow[i\in I]{w^*} m$ shows that m is the integration against χ_Ω, i.e., $m = \nu$. □

Proposition 5.33 (Margulis, Sullivan). *For $n \geq 5$, the group $SO_n(\mathbb{R})$ contains a dense finitely generated subgroup Γ which has property (T).*

Their approach is based on constructing arithmetic lattices in semi-simple Lie groups with property (T). The construction of the mentioned lattices allows them to be embedded into $SO_n(\mathbb{R})$ densely. Since property (T) is inherited by lattices, this procedure provides the required dense subgroups of $SO_n(\mathbb{R})$ if $n \geq 5$. On the other hand, Zimmer showed that for $n = 3, 4$ $SO_n(\mathbb{R})$ does not contain proper subgroups with property (T). This illustrates that a different tool is needed to solve the Banach-Ruziewich problem in these cases. This different approach comes from the theory of automorphic forms. Drinfeld proved the following result using estimates for Fourier coefficients of certain modular forms, namely the Ramanujan conjecture.

Theorem 5.34 (Drinfeld). *For $n = 3, 4$, the group $SO_n(\mathbb{R})$ contains a dense subgroup Γ such that the associated representation π_λ^0 of Γ on $L_0^2(S^{n-1}, \lambda)$ does not contain the trivial representation weakly.*

Theorem 5.35 (Banach-Ruziewicz problem). *For $n \geq 2$, the Lebesgue measure is the unique rotation-invariant, finitely additive normalised measure defined on all Lebesgue measurable subsets of S^n.*

Proof. Let ν be a finitely additive measure defined on all Lebesgue measurable subsets of S^n invariant under the action of $SO_{n+1}(\mathbb{R})$. This ν is necessarily absolutely continuous with respect to λ. The function

$$m : L^\infty(S^n) \to \mathbb{R} \text{ defined by } m(f) = \int_{S^n} f(g) \,\mathrm{d}\nu(g)$$

is an invariant mean. For $n \geq 5$ $SO_{n+1}(\mathbb{R})$ has a finitely generated dense subgroup Γ with property (T). If $f \in L_0^2(S^n, \lambda)$ is a Γ-invariant function, then f is $SO_{n+1}(\mathbb{R})$-invariant by the density of Γ. Since $f \in L_0^2(S^n, \lambda)$, $f \equiv 0$. This means that the action of Γ on S^n is ergodic. This is equivalent to the fact that π_λ^0 has no non-zero invariant function in $L_0^2(S^n, \lambda)$. Since Γ has property (T), π_λ^0 cannot contain π_0 weakly. The conclusion $m = \lambda$ follows from the previous propositions. □

5.2.3 Kazhdan's property (τ) and expander graphs

In this last section, we present a deep result which links the expander graphs to Kazhdan's property of certain groups. This theorem states that most of the magnificent properties we investigated so far, are the reflections of a relative Kazhdan's property in different mathematical structures creating a strong connection between representation theory, expander graph theory, spectral theory of Riemannian manifolds and measure theory of locally compact groups.

Definition 5.36 (Property (τ))**.** Let Γ be a finitely generated group and $\mathcal{L} = (N_i)_{i\in I}$ a family of finite index normal subgroups of Γ. Let $\mathcal{R} = \{\varphi \in \hat{\Gamma} \mid \exists i\ N_i \subset \operatorname{Ker}\varphi\}$. Γ has property (τ) with respect to the family $\mathcal{L}$ if the trivial representation is isolated in the set $\mathcal{R}$.

Theorem 5.37 (Characterization of Kazhdan's (τ) property)**.** *Let Γ be a finitely generated group generated by a finite symmetric set of generators S. Let $\mathcal{L} = \{N_i\}_{i\in I}$ be a family of finite index normal subgroups of Γ. Then the following conditions are equivalent:*

(i) *Γ has propertry (τ) with respect to $\mathcal{L}$, i.e. there exists an $\varepsilon_1 > 0$ such that if (H, ρ) is a non-trivial unitary irreducible representation of Γ whose kernel contains N_i for some i, then for every $v \in H$ with $\|v\| = 1$, there exits $s \in S$ such that $\|\rho(s)v - v\| \geq \varepsilon_1$.*

(ii) *There exists $\varepsilon_2 > 0$ such that all the Cayley graphs $G_i = \mathcal{G}(\Gamma/N_i, S)$ are expanders with expansion constant $h(G_i) \geq \varepsilon_2$.*

(iii) *There exists $\varepsilon_3 > 0$ such that for all Cayley graphs $G_i = \mathcal{G}(\Gamma/N_i, S)$ $\lambda_1(G_i) \geq \varepsilon_3$.*

If $\Gamma = \pi_1(M)$ for some compact Riemannian manifold M, and M_i are the finite sheeted coverings corresponding to the finite index subgroups N_i, then the above conditions are also equivalent to each of the following:

(iv) *There exists $\varepsilon_4 > 0$ such that $h(M_i) \geq \varepsilon_4$ for all $i \in I$.*

(v) *There exists $\varepsilon_5 > 0$ such that $\lambda_1(M_i) \geq \varepsilon_5$ for all $i \in I$.*

Let $G = \varprojlim(\Gamma/N_i)$ be the profinite completion of Γ with respect to $\mathcal{L}$, then the above conditions are also equivalent to the following:

(vi) *The Haar measure on G is the only Γ-invariant mean on G.*

A good introduction to the theory of Kazhdan property is [DlHV08]. The results presented in this chapter and the proofs also can be found in [Lub10].

Bibliography

[BB13] Valentin Blomer and Farrell Brumley. The role of the Ramanujan conjecture in analytic number theory. *Bulletin of the American Mathematical Society*, 50(2):267–320, 2013.

[Bum98] Daniel Bump. *Automorphic forms and representations.* Cambridge University Press, 1998.

[Bus10] Peter Buser. *Geometry and spectra of compact Riemann surfaces.* Springer Science & Business Media, 2010.

[DlHV08] Pierre De la Harpe and Bachir Bekka Valette, Alain. *Kazhdan's Property (T). New Mathematical Monographs.* Cambridge University Press, 2008.

[DS05] Fred Diamond and Jerry Michael Shurman. *A first course in modular forms.* Springer, 2005.

[DSV03] Giuliana Davidoff, Peter Sarnak, and Alain Valette. *Elementary number theory, group theory and Ramanujan graphs.* Cambridge University Press, 2003.

[FK92] Hershel M Farkas and Irwin Kra. Riemann surfaces. In *Riemann surfaces*, pages 9–31. Springer, 1992.

[Fol16] Gerald B Folland. *A course in abstract harmonic analysis.* CRC press, 2016.

[Fri91] Joel Friedman. The spectra of infinite hypertrees. *SIAM Journal on Computing*, 20(5):951–961, 1991.

[Hal12] Gábor Halász. *Lecture notes: Riemann felületek.* 2012.

[HLW06] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006.

[Iwa95] Henryk Iwaniec. *Spectral methods of automorphic forms.* American Mathematical Soc., 1995.

[Iwa97] Henryk Iwaniec. *Topics in classical automorphic forms.* American Mathematical Soc., 1997.

[Kob12] Neal I Koblitz. *Introduction to elliptic curves and modular forms.* Springer Science & Business Media, 2012.

[Kow17] Emmanuel Kowalski. *Lecture notes: Introduction to expander graphs.* 2017.

[Lab15] Olivier Lablée. *Spectral Theory in Riemannian Geometry.* 2015.

[Led94] Michel Ledoux. A simple analytic proof of an inequality by P. Buser. *Proceedings of the American Mathematical Society*, 121(3):951–959, 1994.

[Li93] Wen-Ching Winnie Li. A survey of Ramanujan graphs. *Arithmetic, Geometry, and Coding Theory, Luminy, France*, pages 127–143, 1993.

[LPS86] Alexander Lubotzky, Ralph Phillips, and Peter Sarnak. Explicit expanders and the Ramanujan conjectures. In *Proceedings of the eighteenth annual ACM symposium on Theory of computing*, pages 240–246. ACM, 1986.

[LPS88] Alexander Lubotzky, Ralph Phillips, and Peter Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

[Lub94] Alexander Lubotzky. Banach-Ruziewicz problem for $n = 2, 3$; Ramanujan graphs. In *Discrete Groups, Expanding Graphs and Invariant Measures*, pages 85–100. Springer, 1994.

[Lub10] Alexander Lubotzky. *Discrete groups, expanding graphs and invariant measures.* Springer Science & Business Media, 2010.

[Lub12] Alexander Lubotzky. Expander graphs in pure and applied mathematics. *Bulletin of the American Mathematical Society*, 49(1):113–162, 2012.

[Miy06] Toshitsune Miyake. *Modular forms.* Springer Science & Business Media, 2006.

[Mur03] M Ram Murty. Ramanujan graphs. *Journal of the Ramanujan Mathematical Society*, 18(1):33–52, 2003.

[Pá16] Péter Pál Pálfy. *Lecture notes: Csoportelmélet.* 2016. Version 3.

[Ram16] Srinivasa Ramanujan. On certain arithmetical functions. *Trans. Cambridge Philos. Soc*, 22(9):159–184, 1916.

[Ros97] Steven Rosenberg. *The Laplacian on a Riemannian manifold: an introduction to analysis on manifolds.* Cambridge University Press, 1997.

[Sar90] Peter Sarnak. *Some applications of modular forms.* Cambridge University Press, 1990.

[SS03] Elias M Stein and Rami Shakarchi. Complex analysis. Princeton lectures in analysis, ii, Princeton University Press, 2003.

[Sul81] Dennis Sullivan. For $n > 3$ there is only one finitely additive rotationally invariant measure on the n-sphere defined on all lebesgue measurable subsets. *Bulletin (New Series) of the American Mathematical Society*, 4(1):121–123, 1981.

[Sun88] Toshikazu Sunada. Fundamental groups and Laplacians. In *Geometry and analysis on manifolds*, pages 248–277. Springer, 1988.

[Ter10] Audrey Terras. *Zeta functions of graphs: a stroll through the garden.* Cambridge University Press, 2010.

[Zá17] Gergely Zábrádi. *Lecture notes: Algebrai számelmélet.* 2017.